

Economic revolution postponed for now

Governments have been succeeding in hiding from the consequences of industrial change. When will they begin to prepare for what lies ahead?

ODD things are happening to the world's economy, with the oddest consequences, one of which is that economists, the practitioners of the "dismal science", are now more dismal than usual. The familiar shrugging of the shoulders by which they disclaim the accuracy of their predictions is now often accompanied by protestations that they no longer understand the basis on which predictions should be made. Critics should be charitable. The conundrums which now abound betoken structural changes in the economic pattern of the world whose full effects are far from being realized.

The symptoms of impending change are striking. Here are a few. Earlier this year, it was widely expected that the dramatic decrease by about two-thirds of the international price of oil would have the effect of stimulating the economy of the industrialized West; now, after half a year, it seems more likely that economic activity will decline, which is one of the reasons why the New York Federal Reserve board, the US central bank, last week reduced the general level of interest rates to 5.5 per cent. During the same period, the value of the US dollar on the foreign exchanges has fallen, against some currencies by as much as a quarter, but there is no sign as yet that the US trade deficit, obdurately stuck above \$100,000 million a year, will be abated in the near future.

Meanwhile, the huge amount of debt run-up by the developing countries of the world with the commercial banks of the West remains substantially unchanged; after three years of fear that defaults might bring the whole house of cards crashing down, there has evolved a network of arrangements by which the debts persist (interest continues to be paid) but not even the banking community knows when, if ever, the principal will be repaid. Nevertheless, the stock markets of the world have just sustained a year-long bull market, bidding up the value of industrial stocks and shares to unprecedented levels. The proportions of people out of work in advanced societies are also unprecedented, as are the low prices of the commodities (not only oil) by which the developing countries of the world must hope to earn a living. What can be going on?

Paradox

The most sensible reading of these paradoxical events is that they mark the emergence of structural change in the pattern of economic activity on a scale that financial institutions can contemplate only in disbelief. The case of economic relations between the United States and Japan is a familiar pointer to what is happening. These years, the trade surplus of Japan amounts to roughly half of the deficit of the United States, but the effects of this imbalance are concealed by the willingness of Japanese to spend their surplus on investment in the United States, where steadily increasing proportions of industrial and commercial assets are owned by foreigners. The result is that the dollar has not depreciated against the yen by nearly as much as it would have done in normal circumstances. Japan does not seem as rich as it really is, while the true poverty of the United States is temporarily concealed. Both governments appear to benefit from these arrangements; Japan would no more welcome a more rapid increase of personal prosperity than the United

States would enjoy a reduction of living standards. But this trend can be concealed only for a time.

Much the same is true of the stratagems by which governments are seeking to conceal from themselves the patterns of impending economic change. Both in the United States and Western Europe (but also in Japan), governments spend substantial proportions of their domestic revenues on the support of agriculture, for example. In the United States, the cost will exceed \$30,000 million in the year ahead, something like a sixth of last year's budget deficit and rather more in 1987 if the US Congress manages to keep the total budget within the limits of the Gramm-Rudman Act.

The European Community's Common Agricultural Policy has the same effect. For the governments concerned, it seems preferable to sustain high-cost food production at levels the market cannot absorb (both by direct subsidy and by devices such as the artificial restraint of farm productivity such as the nonsensical European decision two weeks ago to ban the use of animal hormones — see page 762). The losers are not merely taxpayers but countries elsewhere in the world, chiefly developing countries, that might look for a more prosperous future in the development of trade in food.

Hiding

Once-prosperous governments are also bent on concealing from themselves the consequences of both the shift from manufacturing to service industries, under way for the past thirty years, and the shift, within manufacturing, from the products of heavy industry (such as steel) to those of light precision industry (such as electronics). Part of the motivation for this unwillingness to face reality is the sheer discomfort of rapid change, but there is also an underlying fear that countries such as Britain that used to be successful at heavy engineering are ill-equipped with the skills needed for survival in what has become the modern world. One glaring sign of the general fearfulness is the attempt now being made, by an alliance of European and North American electronics manufacturers and music publishers, to persuade Japanese electronics manufacturers to withhold from the market the digital audio-recording equipment they are bursting to sell. If this scheme succeeds, where will protectionism end? Persistently high unemployment, made socially just acceptable by welfare schemes (as it should be), is a mark of the mismatch between the pattern of supply and demand.

This state of affairs cannot persist indefinitely. Water will sooner or later find its own level. But the readjustment, when it comes, could be catastrophic. What will happen, for example, in the slow-moving industrial communities when it turns out that the industrial equities whose prices have been bid up over the past several months by financial institutions buying financial assets with the intention of paying the huge impending cost of people's pensions turn out to be much less valuable than now? In this period of change suspended, the more far-sighted governments should be preparing themselves for what lies ahead, chiefly by investing in the intellectual skills that will be required. For most of them, unfortunately, propping up the past is the easier choice. □



EPA discovers radon

The US Environmental Protection Agency risks stirring up a hornet's nest over radon.

It is a great shame that even the natural life is not entirely free from hazard, but that appears to be the case. For centuries, no doubt millennia, people and proto-people have been dying prematurely of natural causes — catastrophes, infections, climatic fluctuations and starvation, for example. Much of what passes for modern civilization is a marvellous defence against hazards such as these. Housing, by keeping people warm and dry, helps powerfully in this sense, which it is why it is one of life's little ironies that housing has also emerged, in the past few decades only, as a hazard in its own right. For houses, especially when they are of mineral construction such as stone, are also either a source of radon, the radioactive rare gas whose atoms are inescapably intermediates in the radioactive decay of the uranium and thorium series, or traps for it. One consequence is that all creatures on the surface of the Earth take in radon with every breath they breathe. Another is that those who live in houses may be exposed to extra amounts of radon generated from the construction materials. A third is that, in fashionable regard for energy conservation, houses are now less well ventilated and thus more efficient traps for radon.

The ubiquity of radon has been known for almost as long as radioactivity itself, the best part of a century. Only in comparatively recent decades, with better measurement techniques and accompanying conceptual refinements, has it become apparent that radon exposure must be, for most people, a greater source of biological hazard than other natural sources of radiation. The 1984 report of the UN Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) gives a good synoptic account of the problem. The hazards of this exposure, chiefly the risk of usually fatal lung cancer, have been estimated from records of the incidence of lung cancer in people occupationally exposed to radon in, for example, uranium mines (see, for example, Evans *et al. Nature* **290**, 98; 1981). In very round numbers, the radon contributes more than a half of the average effective dose equivalent of about 2.0 mSv a year which people acquire on account of exposure to natural sources of radiation. The dose due to cosmic rays is roughly 0.30 mSv a year, varying with latitude and altitude.

Most public authorities appear to have awakened only late in the day to the importance of radon as a natural hazard, but the US Environmental Protection Agency (EPA) has been later than most. It seems now to be embarking on a crusade against radon pollution with a zeal and innocence that suggest it may have heard of radon only recently. In Washington two weeks ago, James A. Barnes, deputy administrator of EPA, acknowledged that "radon is a serious public health problem" and that "virtually no one" recognized it as such eighteen months ago. He added somewhat regretfully that radon, being a natural hazard, "does not lend itself to traditional regulatory solutions".

Precautions

What EPA has done so far is to publish two helpful booklets, one for members of the public explaining what the problem is and the other for householders and/or their builders suggesting what might be done about it (sealing off ground beneath houses, better ventilation and so on). There is also a modest programme of survey and demonstration, in conjunction with the states (especially Florida, Pennsylvania, New Jersey and New York), and the promise of more action in the future.

How to give radon the attention it deserves without creating panic? Part of the difficulty is that the two isotopes, ^{220}Rn (commonly called thoron) and ^{222}Rn , with half-lives of 55 seconds and 3.8 days respectively, are ubiquitous only in the sense that small concentrations are found in the atmosphere everywhere; on the basis of the total content of radon in the atmosphere (which,

measured in becquerels, is merely an order of magnitude less than that released from the Chernobyl reactor in April), the activity at ground-level should range from 1 to 100 Bq m⁻³, depending on the weather. But there are huge variations from one place to another, depending on such things as the upward flux of radon from the soil in the immediate locality, the flux from buildings and volcanic regions and even that artificially released into the atmosphere by drilling holes through granitic rocks in the pursuit of petroleum or of geothermal steam.

This is why the radon problem is a haunting problem. While the average effect of atmospheric radon on people living in the United States may be small, enough people are probably subjected to enough radiation dose for the personal consequences to be significant. Mr Barnes said that there may be 8 million people whose involuntary risk of death from lung (and other) cancer caused by radon and its decay products is comparable with that smoking from a pack of cigarettes each day, a lifetime chance of, say, less than one in six.

Costs

Mr Barnes' problem (and other people's) is that there is no way of telling what will be the financial consequences of avoidance. The retro-construction of buildings is not cheap, especially when the few people competent to do the work know that there is a panic on. So, in a roundabout way, the radon question raises a more general and difficult conundrum: by what means does a civilized community strike a balance between precaution and cost in the pursuit of an environment free from a natural risk? Can the community afford the cost of being as free from avoidable risk as it would like to be?

In due course, no doubt, in the special circumstances of the United States, the market may help solve Mr Barnes' problem. Realtors may be required to quote measurements of radon doses when they offer houses for sale, and may find that higher doses mean cheaper prices; that exceptional dwelling will be significantly if stochastically hazardous, people will for the first time find themselves picking and choosing where to live on grounds of radiation exposure.

Two particular considerations complicate EPA's problem: houses built in the 1950s in western states with US funds from materials salvaged from the tailings of uranium mines and subsequently found to leak radon were dealt with stringently at government expense, while the more recent recognition that radon concentrations in the so-called Reading triangle reaching north from Pennsylvania affect millions of people, on whose behalf such stringent standards will not be as readily applied.

Buildings accentuate the natural patchiness of radon distribution. The materials of which they are constructed are a source of radon, while the fact that they are of necessity enclosed allows them to become storage reservoirs for radon from the subsoil. The concentration of radon in the atmosphere of a room will be determined by competition between the rate at which the radon accumulates from external sources, abated by its natural decay, and ventilation in the sense of the rate at which air is exchanged with less contaminated air from outside.

The consequence is that radon radiation doses to the tissues of the human lung vary enormously with location and circumstances. The mean dose (arithmetical, geometrical or some other) will reflect regional characteristics, but the doses to which individuals are exposed will more often be very much greater than very much smaller (which is to say that the distribution is skewed). EPA's problem is that of helping individuals especially at risk to identify themselves, and perhaps to protect themselves, without alarming others. That, no doubt, is why the agency sounds so much like a newcomer to the scene. But it must also face the more enduring and, in the long run, more daunting task of knowing what to say to people whose radiation exposure is about average for where they live, but who know (perhaps by reading journals such as this) that their norm is higher than that of others. □

Research donations

Tax bill provides fewer breaks for US education

Washington

CHEERS rang out late Saturday night 17 August in a room crowded with lobbyists, journalists and legislators, as a House of Representatives and Senate conference committee reached an agreement on a major overhaul of US tax law. The agreement came only after weeks of give and take among conferees, watched and prodded by a dizzying array of lobbyists, as they struggled to find a bill acceptable to both houses of Congress. The new law has the White House's blessing, and final passage is expected when legislators return from their summer recess.

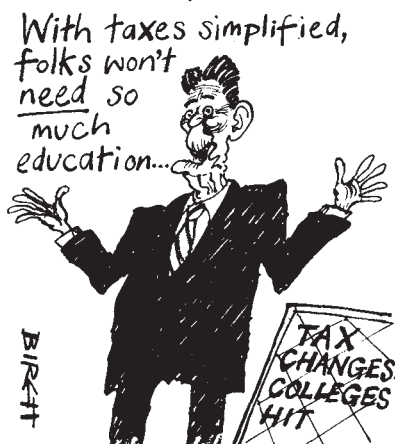
The new tax law will mean significant changes for all sectors of the economy. The number of tax brackets for individual taxpayers will be reduced to two, with a maximum rate of 28 per cent, down from the current 50 per cent. For corporations, the top rate drops from 46 to 34 per cent. The bill is not intended to raise or lower total revenues generated by taxes. Instead, it is an attempt at a more equitable system, with narrower loopholes and fewer breaks for special interest groups. The tax burden will fall more heavily on businesses under the new law, although lowered rates for individuals will be offset by the loss of certain allowances.

Few segments of the economy emerged as clear-cut winners or losers, but one self-declared loser is higher education. Thomas Head, senior federal relations officer for the Association of American Universities (AAU), says education took a beating on nearly every item on its tax agenda. Head feels the biggest blow is the way gifts of appreciated property are treated. Under current law, the entire value of a donated gift can be deducted from taxable income. In the new law, the amount the gift has appreciated is used in calculating a special "alternative minimum" tax that is usually applied to wealthier taxpayers — those most likely to make sizeable donations. The cost of giving appreciated property is therefore much increased. Sheldon Steinbach, general counsel for the American Council on Education, says that 40 per cent of gifts over \$5,000 to universities and colleges come in the form of appreciated property.

Gifts from smaller donors are also likely to decrease because of a change in the reporting requirements for charitable contributions. Taxpayers are offered a choice between taking a standard deduction and itemizing their expenses; under the new law, only itemizers may claim a deduction for their charitable contributions. Only 20

per cent of tax payers are expected to itemize their deductions under the new law. Also limiting donations will be the lowered tax rates, as tax savings from contributions will be proportionately lower.

Gifts to charitable institutions, including those that support scientific research, will also decrease markedly under the new law. Dr Lawrence Lindsey of Harvard University has used an econo-



metric model to predict that total charitable donations in fiscal year 1988 would have been between \$72,000 million and \$75,000 million. But with the new rules total donations will fall by about 14 per cent, and donations by the wealthiest taxpayers by about 36 per cent, or \$1,000 million. The drop in charitable contributions that will be brought about by the inability of non-itemizers to deduct contributions is estimated at \$6,000 million, and that due to the overall drop in tax rates at \$4,000 million. Some charities are planning major fund-raising campaigns this year to take advantage of the existing law before it is superseded.

Major private research universities expect to be hurt by restrictions on access to the tax-free bond market. Tax-free bonds generally have lower interest rates than taxable ones, making them a cheaper way to borrow money. The new law places a \$150 million cap per institution on these bonds. Says David Morse, director of federal relations at the University of Pennsylvania, as many as twenty large universities will be shut out of the tax-free bond market. Smaller private colleges will be unaffected.

Universities are also dismayed by the change in tax status for scholarships and fellowships given to students. In the past, these funds have not been subject to income tax, but the new law makes any funds not applied directly to tuition or re-

quired equipment taxable. AAU's Head points out that since gifts are not taxable, students whose families can afford to give them money for their education would pay no taxes, but a poorer student on a scholarship could face a tax liability.

Another potentially expensive change for universities and colleges is the treatment of pension plans, says Head. These must be shown to apply equally to all university employees. Head says that most plans tend to be more favourable to faculty, and universities may have to cut back on faculty benefits to make plans equitable. In any case proving that plans are equitable will be a drain on university resources.

Not all the news for universities is bad. A new 20 per cent tax credit for industry support of basic research performed at universities and tax-exempt, independent research institutes appears in the new tax law. Dr. Melvin Eggers, Chancellor of the Syracuse University and vice-chairman of the Coalition for the Advancement of Industrial Technology (CAIT), an *ad hoc* group backing the new tax credit, says it will encourage closer cooperation between company and university researchers, speeding up commercialization of research efforts. According to CAIT's estimates, the new credit could bring as much as \$475 million in new research investments at universities. This is a relatively small amount compared to the estimated \$50,000 million industry spends annually on research and development, but supporters believe it is a step in the right direction. The credit will be available through 1988.

Despite criticism that has been levelled against it in the past (see *Nature* 310, 615; 1984), the existing research and development tax credit is preserved in the new tax bill through 1988, although the rate drops from 25 to 20 per cent. Robert Lawrence, a senior fellow at the Brookings Institution, says that in real terms the credit only amounts to about 6 per cent because of the way it is calculated — the credit applies to spending above the previous three years' average. Only 65 per cent of money spent in supporting university-based research figures in the credit calculations. While some have called for eliminating the research and development tax credit, Lawrence thinks the government must give industry even stronger incentives to encourage more spending on research.

The new tax law's chances for passage when Congress returns are good, but some changes may still occur. The rationale for the tax reforms was to deemphasize taxation as a driving force behind business decisions. But taxes are bound to remain an important consideration for corporate decisions, and whole legions of tax lawyers will be kept busy figuring out just what the new realities are.

Joseph Palca & Tim Beardsley

European agriculture

Politics before scientific advice

LAST week's decision by the British agriculture minister to ban the use of growth-promoting hormones in cattle, following a European Commission directive to that effect, has drawn attention to an extraordinary incident last October. Then, the European Commissioner for Agriculture, Mr Frans Andriessen, suddenly suspended his scientific committee just four days before it was to prepare its final advice on the use of the hormones.

The committee, chaired by Professor G. E. Lamming of the Department of Physiology and Environmental Studies at the University of Nottingham, had reported in 1983 that the use of natural hormones such as oestradiol, progesterone, and testosterone in the form of implants in the ear of an animal was harmless to final consumers of the meat, even to pre-pubescent children, potentially the most sensitive group. Last year the committee, containing 22 senior European endocrinologists and toxicologists, was putting the final touches to its report on 'zenobiotics', or artificial hormones, such as zeronol and trenbolone when Commissioner Andriessen gave the order to stop. This effectively gagged the scientists as they are forbidden to reveal the results of their five-year study without the permission of the Commission; but it is widely believed that they would have given zenobiotics the same clean bill of health, provided use were controlled and monitored, as that gained by natural hormones. Just two months later, Andriessen successfully steered through the Council of Ministers a directive banning all agricultural use of hormones.

The British government, facing less public opposition to hormone use than its continental partners, voted against the ban in the Council of Ministers. But, faced with a possible loss of European markets if Britain failed to ban hormones, it now has little choice but to implement the directive. According to the Ministry of Agriculture, Fisheries and Food (MAFF), British meat exports to Europe are worth some £400 million annually. The use of hormones (to make the 50 per cent of British beef which comes from castrated cattle grow faster and produce leaner meat) probably adds only £40 million to meat values, MAFF estimates. Nevertheless Britain is pursuing its action in the European Court challenging the legality of the vote on the directive last December.

Lamming says he "does not want to get involved in politics", and is bound by contract to say nothing about his committee's report on the zenobiotics, but he is clearly disturbed at the possible consequences of the blanket ban on hormone growth promoters now in place in Europe. He believes that the ban could make the situation

more, rather than less, dangerous. Hormones in open use, are supplied from slow-release implants in the ear, a part of the animal not used for meat and where the implant can easily be detected. The implants diffuse small amount of hormones through the body of the animal; a pre-pubescent girl having 100 times more testosterone in her body than she could receive from a normal diet of hormone-treated meat; or a pre-pubescent boy 300 times the oestradiol he might receive from the same source. With the hormones banned, unscrupulous farmers will bury

the implants deep in the muscle tissue of the animals, probably in the most expensive cuts where inspectors are least likely to disturb a carcass, so certain slices of meat could contain far above average levels of hormone.

Lamming and his committee are, to say the least, put out by the Commission's decision. "We've done five years' work, collected voluminous data, and can't publish". Lamming for one will not be joining an *ad hoc* European committee again. But the major issue, which has risen again and again in Europe, is the preparedness of European ministers, who after all backed Andriessen's directive, to go ahead without listening to effective scientific advice.

Robert Walgate

Japanese space science

Joint project for solar maximum

Tokyo

JAPAN's space programme may be tiny compared with that of the United States but its size is not preventing it offering a little hospitality to others. The Institute of Space and Astronautical Science (ISAS) has invited US researchers to place a soft X-ray telescope aboard their solar observation satellite scheduled for launch at the start of the next solar maximum in 1991.

The Japanese High Energy Solar Physics (HESP) mission follows ISAS's Hinotori satellite and the US Solar Maximum Mission that observed solar flares during the last solar maximum in 1981-82. The new ISAS solar satellite, Solar A, will be placed in low-Earth orbit where the US soft X-ray telescope and a Japanese hard X-ray telescope will be trained on the Sun.

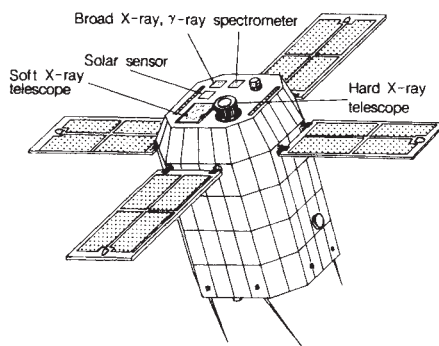
Designed to give high spatial and time resolution, both telescopes will allow second-by-second image analysis of solar flares as they grow and die. The hard X-rays will reveal the hottest (greater than 30 million kelvin) kernel of flares during their flash phase, whereas the soft X-rays will provide images of the cooler pre- and post-flash phases. If there is room for them, spectrometers will also be carried aboard the satellite.

According to the director of ISAS, Professor Minoru Oda, US participation in HESP will be funded through NASA's Explorer programme. An announcement of opportunity was made by NASA last March and official selection of the US team is expected this month.

Difficulties in setting up the joint project occurred, according to Oda, because of differences in approach between the United States and Japan. Whereas ISAS scientists gradually develop, refine and polish a project, and obtain final budget approval once a fairly concrete plan has emerged, NASA starts with a concrete plan and budget and solicits proposals. Other problems arose over the design

nation of principal investigators. But eventually, says Oda, a *tamamushi* agreement was reached (the *tamamushi* is an iridescent beetle whose colour changes according to the angle you look at it).

The satellite is expected to be about 2 m long and 1 m wide with three-axis stabilization to point the long axis towards the Sun. Weighing about 400 kg it will be launched to an altitude of 550-600 km by a



US and Japanese X-ray telescopes will share the new satellite.

Mu-3S-II rocket from the Kagoshima Space Centre during the August-September 1991 launch window. Apart from the US telescope, the total budget for the mission, which awaits final approval by Japan's Ministry of Finance, is expected to be about 4,000 million yen (£17 million), a mere fraction of the cost of launching a comparable satellite in the United States.

The cheapness of ISAS missions is easy to explain. ISAS scientists design, test, launch, monitor and sometimes even build their satellites and rockets, and although work is farmed out to private companies such as Nissan and NEC, there is no principal contractor. Final liability thus rests with ISAS, which substantially reduces the costs. Such an approach also has the advantage of great flexibility, but as Oda points out, it leaves precious little time for science.

David Swinbanks

Chernobyl report

Drama of human perversity

Vienna

THE world's nuclear engineers gathered to hear the Soviet account of the Chernobyl accident on 26 April do not yet know whether they are at a wake for nuclear power or at a launching ceremony for a new generation of safe reactors. Nervously, they have crossed their fingers and praised the Soviets for their frankness.

The head of the Soviet delegation, Academician V.A. Lugasov, won over many critics with a gripping five-hour presentation of the authentic accident report, including personal asides hinting at disagreements on technical matters among Soviet officials. The highlight of the presentation was a video film, made under obviously hazardous conditions, showing the huge scale of the damage caused by the explosion of the No.4 reactor, with its contiguous neighbour miraculously undamaged. As in films of the *Titanic* at the bottom of the Atlantic, four of the eight reactor cooling pumps picked out through the gloom by a flickering spotlight seem as pristine as when they left the factory.

By Lugasov's account, the causes of the accident were not so much human error as human perversity. Plant operators in a hurry to complete a test of the reactor turbine's capacity to span a few seconds interruption of power supply, itself "poorly" designed, dealt with successive danger signals by disabling the safety systems meant to deal with them. In resignation, Lugasov said that when the RMBK-1000 reactor was designed 20 years ago, people had thought that human beings would be more reliable than automatic safety systems. Now he was not so sure.

The Soviet delegation has nevertheless acknowledged several technical defects of the reactor, whose merits are said to include simplicity and low cost. Lugasov pointed in particular to the disadvantage that, as excess steam forms in the water-cooled fuel channels, the reactivity of the reactor increases and also the complexity of the control system required to keep the power output stable.

Other RMBK reactors operating in the Soviet Union will be modified by 1987, chiefly by increasing the fuel enrichment from 2.0 to 2.4 per cent, with an estimated loss of 10 per cent of power output. Thereafter, there will be more radical modifications not yet decided.

The future of the Chernobyl site is much less clear. Two of the four reactors completed at Chernobyl have been decontaminated, but cannot yet be brought back into action for lack of housing for the operating staff now that the towns of Pripyat and Chernobyl are uninhabitable. Decisions remain to be made about reactor No.3 and the two other reactors whose

construction was under way at the time of the accident.

Lugasov's statement has also cleared up some of the questions that have been puzzling the West since the end of April. Thus, he said, there was no undue delay before the plant operators told Moscow of the accident, which meant that a team of specialists could be despatched to Kiev by 8 p.m. on 26 April. But the plant operators subsequently transmitted misleading information about the state of the reactor, the effect of which was to underplay the damage to the reactor and the release of radioactivity.

Lugasov also stoutly defended the decision not to evacuate 49,000 people from Pripyat until 11 a.m. on 27 April, more than 30 hours after the accident. "Strange as it may seem", he said, there was no significant contamination until 9 p.m. on 26 April, when it was judged safer to keep people indoors overnight.

The second wave of radioactivity after a week's delay was explained this week by the heating of the damaged reactor core to an estimated 2,000°C because of the thermal insulation of the clay and sand which, with other materials, had by then been dumped on top of it. More could not be added for fear of making the reactor foundation collapse, but there seems to have been no danger of a fuel meltdown.

John Maddox

Private money for AIDS research

Washington

THE US Public Health Service (PHS) wants to encourage private sector involvement in its efforts to find a vaccine to prevent acquired immune deficiency syndrome (AIDS). To that end, PHS spelled out in the 23 August *Federal Register* a more formal framework for collaborative efforts between the government and private institutions. The hope is that such collaborations will speed up the search for an AIDS vaccine, but some wonder whether the government has a sufficiently succulent "carrot" to encourage industry to play along.

Under the framework, PHS will not provide any financial assistance to its collaborators. Instead, it will offer knowhow from its component agencies — the Alcohol, Drug Abuse and Mental Health Administration, the Centers for Disease Control, the Food and Drug Administration (FDA) and the National Institutes of Health (NIH). This assistance could include patent licences, research results, facilities, animal models and animal testing, and help in formulating clinical protocols and clinical trials. PHS is interested in finding collaborators who will be involved in all steps in vaccine development, from basic molecular biology to clinical trials to marketing and commercial production. Companies unable or unwilling to participate in such a large scale project are encouraged to find others in the private sector to join with them in collaborating with the government.

PHS identifies two basic approaches to vaccine development that it wants to pursue. One is the development of a synthetic vaccine, generating viral antigens from whole virus, or using recombinant DNA techniques, chemical synthesis or anti-idiotypic antibodies. The other is development of a live, genetically altered viral vector or attenuated vaccine. PHS will also consider other approaches, such as passive immunization, if they show promise.

A special committee established by the assistant secretary for health of the Department of Health and Human Services will make final decisions on collaboration following a review by the applicable PHS agency. PHS decisions will be based on an applicant's experience and ability to achieve its stated goals.

In the past, some companies have been reluctant to enter into collaborations with the government for fear that their proprietary research could be made public under the Freedom of Information Act. Lowell Harmison, science adviser for PHS, says companies will be protected from unwarranted disclosures as far as the law allows, but otherwise will have to give careful thought to how they word their applications. George Galasso, assistant director for extramural research at NIH, does not think these caveats will be sufficient. Companies that really have a promising vaccine approach "will go it alone", leaving the "also-rans" to seek government help.

Galasso also wonders whether formal collaborations with the government will speed the development of a vaccine. Scientist-to-scientist relationships have worked well so far, says Galasso, and will continue to be available when the need arises. Dino Dina, director of virology for Chiron Corporation, says having FDA help in developing acceptable standards for animal trials and initial clinical trials would be useful. But he feels a more efficient approach would be to make such information available to all comers.

Dina also worries that previous offers of "collaboration" from the government have turned out to be more take than give. But Harmison says the PHS plan is a fresh approach to cooperative arrangements. He hopes that the offer of government facilities and expertise will encourage companies with interesting ideas about vaccine development to pursue them.

Joseph Palca

Space station

Taming the last lawless frontier

Washington

ALTHOUGH it will be several years before astronauts are working together in a multinational space station, it is not too soon for Congress to start considering legal questions that might arise, according to the Office of Technology Assessment (OTA).

Some US laws, Congress's research office notes in a recent background paper*, do not make much sense in space. The US Uniform Commercial Code, for example, with its definitions of personal property and real estate, and what is movable and what immovable, could not be applied "without serious uncertainty". The wisdom of applying the Buy-America Act is similarly unclear. And OTA notes it might be "inappropriate" to apply the Fair Labour Standards Act, with its stipulated 8-hour maximum working day.

More seriously, the application of intellectual property, product liability and export laws in a space station is likely to be important for commercial companies working on the space station. Although the 1967 Outer Space Treaty established some basic principles governing countries' responsibilities and liabilities in space, it leaves many questions unanswered. OTA recommends that Congress start thinking

cult to sort out in space. There is already a variety of opinions on how far US patent law can be applied, and attempts in Congress to clarify the situation have not been very successful. Real disputes are likely to arise, since there are important differences in patent laws between nations.

In the case of criminal law, the question on enforceability also arises. So far crews have been highly disciplined and engaged in specific tasks. But as more people start to live and work in space there may be disputes. On space shuttle missions, the

commander has broad authority to enforce discipline. But if a British astronaut were to assault an American astronaut in the British portion of a space station, US laws could not be enforced without a prior agreement.

One way of avoiding complicated legal entanglements, according to OTA, would be for participating countries to enter in to "pre-launch agreements" modelled on the "status of force agreements" between members of the North Atlantic Treaty Organization which clarify legal questions about forces stationed in different countries.

Tim Beardsley

*Space Stations and the Law: Selected Legal Issues. Office of Technology Assessment Background Paper. 1986.

UK universities

Government yields just a little

THE British government now seems willing to provide the extra money needed to prevent university closures — provided, that is, universities agree to reform and to their performance being monitored.

The news represents a victory for those who believe the only way to persuade the government to invest more heavily in the universities is by showing a quiet willingness to make improvements (or by being submissive, as cynics have it). Beginning in 1981, financial cuts have brought several universities to their knees: at least four are thought likely to have to close during 1987 if help is not forthcoming. The political embarrassment that this could generate in an election year, plus the new willingness of the universities to accept reform and monitoring seems, according to reports from the Department of Education and Science, to have persuaded the Secretary of State, Mr Kenneth Baker, to take out his chequebook, if not actually to sign the cheque.

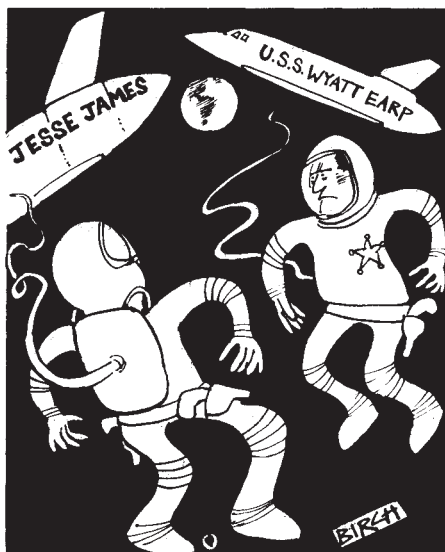
Some of the changes the universities may have to accept are contained in the report on "Academic Standards in Universities" just released from the Committee of Vice Chancellors and Principals. Perhaps its most radical suggestion, for British academics at least, is that students should help appraise courses and lecturers. A host of other course evaluation measures are suggested: the number of applicants, the qualifications of entrants, degree results, percentage of drop-outs and so on. The report also seeks to lay down codes of conduct for external examiners, who will ensure that degrees awarded in similar subjects are comparable in standard in different universities; for postgraduate training and research; and for an appeals procedure for postgraduates who fail to be awarded degrees. The latter codes are intended to help deal with the embarrassingly high drop-out and failure rates among post-graduate students. The responsibilities of a doctoral supervisor

are made explicit for the first time and it is recommended that a statement to establish clear mutual expectations between student and supervisor always be drawn up.

The other issues on which the government wants to see progress are financial management, the creation of performance indicators so that whole universities might be appraised, and the rationalization of university departments that are judged too small or too weak. University performance indicators were set out in the Jarrett report last year: they include, among other things, the success of graduates in obtaining employment; publications by staff and citations; patents, inventions, and consultancies; prizes; and papers given at conferences (see *Nature* 314, 393; 1986).

The University Grants Committee has already set in motion an assessment of individual departments' research quality which seems set to lead to the closure of those judged inefficient. This has proved controversial; not only because of the uncertainties involved in assessment (see *Nature* 322, 219; 1986), but because small departments are emerging as the prime candidates for closure. While many will acknowledge this makes sense in areas where a critical mass of researchers is required to maintain expensive laboratory facilities, it may be a different story when it comes to subjects like the history and philosophy of science. There is a strong argument that the present situation, with sixty or so staff scattered over 25 universities, is the best way to allow many scientists to absorb some history and philosophy of science: to say that it should be concentrated in a few places is to deny its general relevance. Similar arguments over the rationalization of departments of philosophy are now provoking public debate and it is too soon to say whether philosophy will be seen as just a discipline for specialists.

Alun Anderson



immediately about how far US federal and state law can be applied in space, and how disputes might be resolved.

How to determine the jurisdiction of different countries over different parts of a multinational space station is seen by OTA as central. Both Europe and Japan are expected to contribute habitable modules; although the simplest thing from the US point of view would be to have everything under US law, this "may be politically unacceptable to other space station partners", OTA notes.

Patent law is likely to be especially diffi-

US defence

Congress takes a moderate line

Washington

GROWING impatience with President Reagan's arms control policies has prompted the House of Representatives to take actions on its own. Before leaving for its summer recess, the Democrat-controlled House passed a decidedly "doveish" defence authorization bill and the Republican-controlled Senate also showed signs that it is ready for a more moderate build-up of US military capabilities.

Both houses of Congress have now completed work on versions of the defence authorization bill. Neither house went along with the administration request for \$320,300 million for the Pentagon: the House authorized \$286,000 million, while the Senate agreed to \$295,000 million.

The issue of chemical weapons has proven particularly divisive for Congress. The House agreed by just one vote to delay procurement of binary nerve-gas weapons until 1 October 1987. The Senate continued a string of close votes on chemical weapons and a measure that would have deleted funds for the Bigeye bomb, a new binary nerve-gas bomb that has failed many of its developmental tests, was defeated by one vote.

Although President Reagan has indicated that he no longer intends to abide by the terms of SALT II, the strategic arms treaty negotiated in 1979 but never ratified by Senate, opposition to that decision has been fairly strong in Congress. The House adopted a measure that would force the President to abide by SALT II by prohibiting funds to be spent on any weapons system that would be in violation of the treaty. The Senate would only go as far as requesting a Pentagon report on the consequences of violating the terms of SALT II and other strategic arms control treaties.

Spending on research on a space defence system was cut in half by the House, but authorized for the full \$278 million requested by the administration in the Senate. The Senate bill also includes \$28.5 million to begin procurement of a space defence system. Both houses made deep cuts in the administrations request for funding for the Strategic Defense Initiative.

The Pentagon officially runs out of money at the end of next month. Even if Congress cannot reach a compromise on the two authorization bills, similar language on weapons reduction schemes is included in appropriations bills working their way through Congress. One way or another, there is sure to be a shoot out on arms control before the Pentagon gets any more money.

Joseph Palca

Chernobyl accident

Fallout pattern puzzles Poles

"HOTSPOTS" of radioactive fallout from Chernobyl are puzzling Polish scientists. Dr Zbigniew Jaworowski of the Central Radiological Protection Institute in Warsaw told a Japanese reporter, Mr Fumihiko Yoshida, of the *Asahi Shimbun* recently. The Japanese media have been particularly successful in elucidating details of the Chernobyl disaster and its aftermath — presumably the governments involved respect their unique positions as the only victims of a nuclear bombing. It was on Japanese television that Dr Valerii Legasov, deputy director of the Kurchatov Nuclear Energy Institute of the Soviet Academy of Sciences, revealed that almost all the safety circuits at the Chernobyl power station had been disabled during the unauthorized experiments that led to the disaster.

Dr Jaworowski, who is a leading adviser to the United Nations on nuclear fall-out, is recorded as saying that the hotspots, which measure several tens to several hundred metres across, exhibit levels of radio-activity some ten times higher than the surrounding area. The spots are roughly circular in shape and, in the immediate post-Chernobyl period, it appears their outline could be mapped by field observers with geiger counters. Analysis shows that ruthenium (an element whose name, ironically, derives

from the Latin designation for Ukraine) is the major radioactive material in the hot-spots, although there are a few in which lanthanum or barium isotopes predominate. Although most common in north-east Poland, they could be observed all over the country, Dr Jaworowski reportedly said.

Details of the level of radioactivity of the hot-spots have not been given, but similar ruthenium particles deposited in Sweden have registered levels of 1,000 – 10,000 becquerels per particle.

The Polish scientists, Yoshida reported, have no idea why the hot-spots are so widely distributed, nor why they consist predominantly of ruthenium. An apparently similar phenomenon in the Mahileu rayon of the Byelorussian SSR (well away from the total exclusion zone) has been attributed to a freak rain-shower, but, as the "self-help" advice issued by underground Solidarity makes clear, the one thing the Poles were longing for in the immediate aftermath of the disaster was rain which would, they hoped, wash away the air-borne hazard. Dr F.B. Smith, an expert on air-borne pollution from the UK Meteorological Office at Bracknell, suggests that the radioactive material was deposited by dew, which is sensitive to the temperature differences generated by soil-type and land use.

Vera Rich

Japanese mega-prizes announced

ONE of science's newest — and biggest — prizes has gone to Professor G. Evelyn Hutchinson of Yale University and Professor Nicole M. LeDouarin of the Institut d'Embryologie, Nogent-sur-Marne.

Each will receive the Kyoto Prize medal and ¥45 million (\$300,000) at a ceremony to be held in Japan in November.

The prize was set up last year by Kazuo Inamori, founder of Kyocera Corporation, Japan's most famous and fastest-growing high-technology ceramics company. In just twenty-five years Kyocera has "by the grace of God" grown to produce annual pre-tax profits of ¥53 thousand million. That has made it possible to endow the Inamori Foundation with ¥20,000 million to provide for the award of an annual prize.

The aim of the prize is to encourage balance between scientific and technological development on the one hand and psychological and emotional maturity on the other: equilibrium between the "ying and the yang, the light and the dark" is sought.

Many would say that equilibrium has

been achieved by Professor Hutchinson who is known both as an essayist and a scientist. He did much to establish ecology as a science and was the first to develop the concept of the ecological niche. Born in England in 1903, Professor Hutchinson is now an emeritus professor and still writing and researching.



Prizewinners LeDouarin and Hutchinson — each with \$300,000 to spend.

Professor LeDouarin, born in 1930, is known for the work that led to the creation of quail-chicken chimaeras. The distinctive properties of the tissues of the two animals make it possible to follow previously-unobservable events in development.

Belief in miracles

SIR—With regard to R.J. Berry's question "What to believe about miracles" (*Nature* 322, 321, 1986), I think it more interesting, and more of a scientific question, to ask how one can understand and explain belief in miracles. Belief in miracles is a widespread and very influential psychological phenomenon, underpinning value systems and cosmologies that have shaped the history of the world. The events called "miracles" may or may not have occurred in the form in which they are described, but their existence as a phenomenon of social fact is very real. This deserves serious investigation.

The main problem with most miracles is the ambiguity of testimony. Often they were reported long after the event, at second-hand; even if the reports are first-hand and attested to "by thousands", it is usually one person who wrote the description. So the data are usually shaky.

There are therefore, a variety of possible hypotheses. One is that the event occurred, and that it was due to divine intervention in some form. A second is that the event did not happen at all, that nothing remotely like the event occurred, but that people at a subsequent date constructed mythical events and actions surrounding charismatic figures. This is a virtually universal psychological phenomenon, and indeed there are certain common myths that existed long before their adoption by Christians — the virgin birth, and the death and resurrection of the God, for example, and many religions whose origins are in floodplains have Noah-like heroes in their mythology.

A third hypothesis is that something happened, but there was inadequate information about it, and the ambiguities are resolved in a way that provides both consistency and charisma — such psychological processes are demonstrated for every first-year student in social psychology in laboratory classes.

These explanations do not deny the possibility of miracles, but they do bring into question Berry's distinction between domains of faith and domains of science by asking what parameters of miracles are scientifically interesting. Indeed, Berry's point about reductionism is an important one, and something that the newer sciences and social sciences have been plagued with. Setting limits to one's science is a way of achieving rigour, by defining what is possible for investigation given the methodological resources available. But such rigour becomes *rigor mortis* if the limits set are those of one's actual and present methodological constraints, not those of some potential and future, more sophisticated methodology. Developments in psychology have frequently depended on researchers going "off limits"

and defining a problem as interesting some time before the methods existed to study it; undoubtedly this is true of all sciences. Berry argues that it is reductionist to say, in effect, that because we cannot study them (given present techniques), miracles are impossible; he is, I would argue, equally reductionist in saying that it is not possible to investigate miracles using scientific measures. It depends which science one applies to the problem.

HELEN WEINREICH HASTE

University of Bath,
School of Humanities and Social Sciences,
Claverton Down,
Bath BA2 7AY, UK

What's in a name?

SIR—That (novel carbon cluster C_{60}) which we call Buckminsterfullerene¹ has excited considerable interest while the name we have chosen has stimulated comment². Although many have approved of this name, there have also been criticisms that it is too long (one should consider the IUPAC name), that it is clumsy (which is irrelevant), that nobody has heard of Buckminster Fuller (the name thus has educational value) or that an injustice has been done to almost everyone who has played with symmetrical objects from Plato and Archimedes to Stanley Matthews (the professional footballer).

The molecule we believe we have discovered consists of a sheet network of carbon atoms linked to form a highly resilient hollow sphere. The names of Plato and Archimedes are linked with the regular and semi-regular solids, an association, by the way, which may not be totally justified. The earliest hollow association I have found is in the work of Leonardo da Vinci³.

Buckminster Fuller showed that a hollow, light, strong structure could be constructed out of a network of struts using the known principle of Euler that, for closure, 12 pentagonal configurations must be dispersed among the hexagonal ones. The spherical structures such as the Montreal Expo '67 dome, the Epcot Dome and numerous radomes are constructed to take advantage of the lightness, strength (the strains are evenly distributed) and of course the internal cavity, that such a geodesic dome affords. C_{60} is a geodesic carbon atom network with a very strong inert structure and a large central cavity that can trap other atoms. These three major properties are important for this molecule's behaviour and have direct analogues in the success of Buckminster Fuller's geodesic structures, factors that are not inherent in other names. C_{60} Buckminsterfullerene itself is only one

member of a set of structures (C_{44} – C_{80}) that appear to be closed and, although they must have 12 pentagonal configurations, do not have to have *icosahedral* symmetry. Buckminster Fuller also considered such structures⁴.

H. W. KROTO

School of Chemistry
and Molecular Sciences,
University of Sussex,
Falmer, Brighton BN1 9QJ, UK

1. Kroto, H.W., Heath, J.R., O'Brien, S.C., Curl, R.F. & Smalley, R.E. *Nature* 318, 162 (1985).

2. Stewart, P.J. *Nature* 319, 444 (1986).

3. Fra Luca Pacioli *De Divina Proportione* Ambrosiana, Milan: Diagrams by Leonardo da Vinci.

4. Buckminster Fuller, R. *Inventions* (St Martins Press, New York, 1983).

Spanish universities

SIR—In two recent issues of *Nature* (319, 710 and 322, 198; 1986) Juan A. Subirana and Pedro Puigdomenech have pointed out the danger of "inbreeding" in the Spanish universities, as a consequence of the new recruiting system for academic posts. That this danger is becoming fact is demonstrated by the first examinations carried out under the new system.

In most cases, the process of appointment seems to have been a mere formality, the final decision being taken before the examination. Nearly all those obtaining posts were the so-called "official candidates", who were proposed and supported by the university departments before the vacancy was publicly announced. Such candidates began with at least two out of the five votes of the selection committee, those of the members of the department concerned. Other applicants, even those with scientific qualifications superior to those of the official candidates, were often eliminated on the basis of subjective arguments. Work carried out at foreign institutions was undervalued, while some foreign candidates were eliminated by questions such as "What are the idiosyncracies of local students?"

In those circumstances, it is unlikely that many Spanish graduate students and postdoctoral fellows now working in foreign scientific institutions will ever get positions at Spanish institutions. This will not encourage the return of such people to Spain. As a logical consequence, the standard of scientific research in Spain (already much lower than in other Western European countries) will not be raised in the future. This is a serious deficiency in a law that otherwise has many positive features and that was intended to improve the general level of Spanish research. It is a matter of great urgency to reform this system by a new or additional policy.

JAVIER NAVAL

Chimie des Protéines,
Institut de Recherches sur le Cancer,
BP 8, 94802 Villejuif, France

Law and genetic testing

SIR—Bains observes that genetic screening may have attendant legal ramifications¹. This has certainly been demonstrated in the United States.

The expanding ability of medical science to predict and detect defects before birth may have important applications in clinical medicine. Genetic counselling and prenatal testing may potentially provide valuable information to patients planning families about the likelihood of various defects in their offspring. The development of various genetic screening techniques, at least in the United States, has further spawned the birth of various classes of lawsuits. "Wrongful pregnancy" actions refer to cases where the parents of a child file a claim for the monetary and emotional damages suffered as the result of giving birth to a healthy, albeit unwanted, child². The action may arise where a child's conception was due to the alleged negligent performing of a sterilization procedure. "Wrongful birth" suits are those instituted by parents claiming that they would have avoided conception or terminated the pregnancy had they been advised of the risks of birth defects in their offspring³. "Wrongful life" cases are instituted by the infant and allege that, as the result of the negligence of the defendant health-care provider, birth has occurred⁴. The infant is claiming essentially that the defendant has wrongfully deprived the parents of information, which would have resulted in the child not being born.

The three classes of lawsuits are thus different. Wrongful pregnancy cases typically involve a healthy, although unwanted, child, whereas wrongful birth actions normally involve planned children who are born deformed. Both actions are normally brought by parents. However, the wrongful life action is brought by the infant. Allegations in recent, selected lawsuits have involved the failure to diagnose Down's syndrome during pregnancy⁵, failure to diagnose rubella suffered by the mother during pregnancy resulting in the birth of a rubella baby with severe congenital defects⁶, failure correctly to type and record maternal blood and the later birth of a child with erythroblastosis fatalis⁷, the "wrongful birth" of children suffering from fetal hydantoin syndrome associated with the mother's use of dilantin during pregnancy⁸, negligent performing of a vasectomy⁹ and the negligent performing of tubal ligation¹⁰.

With the increased knowledge in the field of genetic screening, there has been a concomitant recognition by various courts that the appropriate standard of professional care may require the use of available, pertinent prenatal tests and genetic

counselling, particularly for patients at high risk of having children with birth defects. Perhaps it may be helpful to seek definitive legislative guidance concerning the vexing legal questions potentially raised by genetic screening.

LEO UZYCH

103 Canterbury Drive,
Wallingford, Pennsylvania, USA

1. Bains, W. *Nature* **322**, 20 (1986).
2. *Continental Cas. Co. v. Empire Cas. Co.*, 713 P.2d 384 (1985).
3. *Blake v. Cruz*, 698 P.2d 315 (Idaho 1984).
4. *Harbeson v. Parke-Davis, Inc.*, 656 P.2d 483 (1983).
5. *James G. v. Caserta*, 332 S.E.2d 872 (W. Va. 1985).
6. *Garrison v. Foy*, 486 N.E.2d 5 (Ind. App. 3 Dist. 1985).

Models of man

SIR—In the context of the diversion of psychological research funds to cognitive science and computing projects, the argument below must be considered:

There are two flaws in computational models of man: (1) in human beings there can be no certainty what the programs are; (2) in human beings the database is unknown and different for every person. It therefore follows that computational models, being quite unlike the human case, are doomed to failure.

PAUL KLINE

Department of Psychology,
Washington Singer Laboratories,
University of Exeter,
Exeter EX4 4QG, UK

Help for Africa

SIR—Michael Spencer's letter (*Nature* **322**, 10; 1986) prompted me to read again your leading article "Who will pity Africa?" (*Nature* **321**, 548; 1986). Pity and political posturing seldom assist the objective assessment of any problem, least of all the vast and complicated set of problems that afflict Africa today. Spencer highlights some of the external political and economic pressures, whereas your leading article, like the United Nations' statement of early June, drew attention to the contributions to the chaos made by some of the African governments themselves. All these have to be taken into account, and much else besides, not least the vastness and heterogeneity of Africa itself—some 45 separate states and ecological conditions ranging from complete desert to tropical rain-forest, with montane and temperate regions besides.

You call for "a political clearing-house for good ideas that have already contributed to the improvement of Africa's conditions on a small scale... which... could be spread more widely". I hope your call will be heeded, but I would suggest that the emphasis should be not so much on good ideas as on good projects that have proved themselves over a reasonable period of time.

As Spencer says, there has been no

shortage of good ideas, but some of these have proved to be disastrous. They range from simple wells in arid regions which have attracted so many people and their animals that the surrounding land has been destroyed, to multi-million dollar irrigation and farming projects that have impoverished the local farmers and their land.

Most of Africa never was and never will be "a Garden of Eden"—the ecology is not like that—but there are instances, past and present, where sensible efforts, with or without outside help, have led to real improvements in the lives of ordinary people. We need to hear more about these, and less about who is to blame for the present troubles.

RONALD KEAY

38 Birch Grove,
Cobham,
Surrey, KT11 2HR, UK

Darwin's yellow rain

SIR—In the light of the continuing controversy over yellow rain, it might interest your readers to know that Charles Darwin observed and reported a case of yellow rain. The report was quoted in the 18 July 1863 issue of the *Gardeners' Chronicle and Agricultural Gazette*¹. The letter from Darwin reads in part:

A very slight shower, lasting hardly more than a minute, fell here this morning (July 2) about 10 o'clock. My wife gathering some flowers immediately afterwards noticed that the drops of water appeared yellowish, and that the white roses were all spotted and stained. I did not hear of this circumstance till the evening; I then looked at several roses and Syringas and found them much stained in spots. Between the petals of the double white roses there were still drops of the dirty water: and this when put under the microscope showed numerous brown spherical bodies, 1/1000 of an inch in diameter, and covered with short, conical transparent spines.

Darwin goes on to describe additional small particles "only just visible with a quarter-inch object glass". The author of the article, designated only as M.J.B., reports examining rose petals forwarded by Darwin, and observing "multitudes of irregular bodies so minute as to present the Brownian molecular motion". M.J.B. concludes that "it is quite astonishing what a multitude of bodies are carried about by the wind in the form of dust".

Indeed, it is equally astonishing that we should still be arguing about yellow rain 125 years later.

BRUCE E. FLEURY

Science and Engineering Division,
Howard-Tilton Memorial Library,
Tulane University,
New Orleans, Louisiana 70118, USA

1. *Gdnrs' Chron.* No. 29, 18 July, p.675 (1863); reprinted in *The Collected Papers of Charles Darwin* Vol. 2 (ed. Barrett, P.H.) 81-82 (University of Chicago Press, 1977).

Science and faith

SIR—In seeking to refute the beliefs and arguments of creationists, J. Richard Wakefield (*Nature* 320, 392; 1986) makes four unjustified generalizations to which I must object as a Christian and as a scientist.

My first point deals with the limits to science. Wakefield states categorically that nothing is potentially beyond science. This is questionable. The scientific method is empirical and depends on testing one's hypotheses against what happens in the real world. Its conclusions ("scientific laws") are inductive and provisional, and they remain open to revision in the light of new and more comprehensive experience. Scientific knowledge is still partial and incomplete, and many things lie beyond its boundaries. The day may come when science can satisfactorily describe, predict and explain every phenomenon in the cosmos that *can* be subjected to scientific analysis; but the possibility that the universal set of real phenomena extends far beyond the finite subset of scientifically verifiable phenomena will remain and cannot be disproved by science. How can science ever illuminate what lies behind the beauty of form, of poetry, of music, of humour, and of human emotion? Humans are physical, rational, intuitive, emotional, personal, cultural, social, moral and spiritual beings. Science is only one of the forms of knowledge available to us; we are familiar with different kinds of knowledge in art, in love, in politics and in religion. Science is not served by making exaggerated claims on its behalf, any more than is creationism.

My second objection is to Wakefield's view of faith, which he defines as "belief without evidence". St Paul, however, taught that "faith gives *substance* to our hopes", faith is "the *proving* (or *test*) of things not seen" (*Hebrews* 11.1; New English Bible, Revised Standard Version — emphases mine). In the classical tradition of Christian experience and understanding, faith is trust that leads to action; and the outcome of living by faith, in the testimony of countless Christians, amply substantiates and justifies it. As science generally considers it good practice to test hypotheses *before* discarding them as invalid, I suggest that Wakefield and those who sympathize with his views subject mainstream scriptural Christian faith to the same test before they totally discount it.

Third, according to Wakefield, "we all know that science accepts nothing on faith". Science can provide us with statements about degrees of probability and correlation. To speak of the provisional and ever-cautious statements of science as if they were eternal proclamations of certainty and causality is to distort the truth.

As D.H. Koobs points out (*Nature* 319, 172; 1986), to assert today that the Universe and its life forms occurred spontaneously remains "clearly a matter of faith".

Finally, Wakefield claims that "creationists do their best to twist, lie, fabricate, misrepresent all they can of reality to deceive their followers and the lay public". While this may be true of some creationists, I cannot believe it is true of all. It is possible to hold a creationist view of the origin of life and the Universe with sincerity and intellectual integrity. And in the interests of scientific objectivity, it must be pointed out that scientists, too, have sometimes distorted the truth: let us remember Burt's obsessional prejudices about race and IQ which he dishonestly presented as if they were real experimental findings.

VALERIE VENTURA

21 Mountsfield Court,
Hither Green Lane,
London SE13 6RR, UK

Risk analysis

SIR—S.O. Funtowicz and J.R. Ravetz (*Nature* 321, 644; 1986) are, of course, correct in writing that society must recognize that large improbable accidents can and will happen, whether they be collisions of Boeing 747 aircraft over Wembley Stadium during a Cup Final, or the failure of a nuclear power plant.

Their rejection of probabilistic risk analysis (PRA), however, is factually incorrect. It was *not* used in either of the two examples he cites — Three Mile Island or Chernobyl. In 1977, the US Nuclear Regulatory Commission instructed its staff not to use PRAs in the licensing procedures. As a result, nobody calculated for a Babcock-Wilcox reactor until the afternoon of 28 March 1979. If they had, the weakness of the operating procedure would have been quickly recognized and the accident at Three Mile Island avoided. In spite of numerous reports and discussions, both official and unofficial, nobody in the West has seen a PRA for Chernobyl. I do not believe that a complete one exists.

It is the procedure of thinking carefully about accidents, using logically plausible scenarios, that is our best and possibly our only defence against these scenarios occurring. This is what a PRA accomplishes.

Some responsible, competent person must have thought through the problem enough to derive a number with its uncertainty. Whether Funtowicz and Ravetz believe the number is comparatively irrelevant. But they should not discourage the procedure itself, which is almost the only procedure that we have.

RICHARD WILSON

Department of Physics,
Harvard University,
Cambridge, Massachusetts 02138, USA

Pesticide regulation

SIR—I appreciated Kenneth Mellanby's letter in *Nature* (321, 465; 1986) about my objection (*Nature* 320, 391; 1986) to a sentence in Lord Ashby's review of John Sheail's book (*Nature* 318, 21; 1985).

I have admired the British method of pesticide regulation, and Sheail performed a good service by describing the collaboration there between industry and government.

I certainly did not intend to imply that it was Mellanby's British colleagues who resorted to "such unscientific methods as deliberately distorting or omitting all the data that refuted their (anti-DDT) allegations". Instead I had in mind only the work of several notorious US scientists, especially some employees of our federal Fish and Wildlife Service and a few in the California Department of Fish and Game. Their statements were frequently published, without competent peer review, in magazines such as *Science*, *BioScience*, *Scientific American* and various "bird" journals. Specific refutations of their allegations have been published elsewhere by many concerned US scientists, but were ignored by the US Environmental Protection Agency (EPA).

EPA joined the attack, and sought to direct a worldwide condemnation of agricultural chemicals and chemicals used to alleviate insect-borne disease outbreaks, via its Toxic Substances Control Act (TSCA). In response, British Science Attaché Alan Smith was applauded by over 600 specialists at a meeting called by EPA to discuss the act when he stated: "This draft is like the jabberwocky of Lewis Carroll . . . the language of chemistry mixes uneasily with the language of metaphysics and the overlay of legal jargon makes the whole incomprehensible." After calling it an "absurd piece of gobbledegook", Smith suggested that EPA should "not presume to legislate for the Universe and the whole human race" and that "there is a limit to the number of times even the greatest country in the world can afford to appear ridiculous in international affairs" (*Science* 196, 1182-83; 1977).

Although DDT is no longer used in the United States, false claims about it continue to be repeated. The reasons are varied, but are well-known. The same sort of campaign that succeeded in banning DDT has more recently been directed against 2,4,5-T, malathion, DES and many other useful chemicals of very slight hazard to man or the environment. Unless the scientific community actively exposes such false statements as have been made in journals and indicates disapproval of those who deliberately make them, the human race faces a very bleak future.

J. GORDON EDWARDS

San Jose State University,
San Jose, California 95192, USA

New ways with interatomic forces

Attempts to calculate the properties of real materials from those of their constituents have never been outstandingly successful. New, but still empirical, techniques may help.

SINCE the year dot, or thereabouts, all kinds of people have been calculating the properties of matter of various kinds by making assumptions they know to be inadequate about the forces between constituent particles. The process continues, and discontent with the assumptions is, if anything, made more explicit by the availability of computers and the associated computer codes that will allow the treatment of systems large enough to be considered "realistic". But supply tends to grow to meet demand, and the invention of novel schemes for representing interatomic forces is now also flourishing. One such is a way of representing long-range interatomic forces developed by a group based at the University of Trieste (F. Ercolessi, E. Tosatti and M. Parrinello) and applied to the structure of the surface of gold crystals (*Phys. Rev. Lett.* **57**, 719; 1986).

The problem of calculating *ab initio* the properties of a material is indeed intractable. Even that model of the ideal gas in which the atoms are perfectly elastic solid spheres is useful only for as long as the spheres are not so large, relative to their container (again assumed perfectly elastic) that geometrical effects make them clump together in some sense. More realistically, the forces between pairs of atoms in, say, a monatomic gas are better represented by a distance-dependent mutual potential energy leading to forces that are repulsive at short distances (reflecting the solidity of atoms) which may be offset by attractive forces at greater distances. Van der Waals is the name to conjure with; physically, what happens is that random intra-atomic fluctuations giving individual atoms an electrical dipole moment will tend to synchronize with the instantaneous dipole moments of interacting atoms, dragging the two together, which is why gases whose constituents are either polar or highly polarizable molecules depart most conspicuously from the perfect gas laws.

Mischievously, real substances are more cussed than these simple examples allow. Even in gases, three and multi-body forces are bound to be relevant, and properties may depend not only on the positions of atoms or molecules instantaneously, but on their velocities as well. A further complication, especially when the average inter-particle distance is small, is that there may be literally no limit to the distance over which inter-particle forces

exert themselves because of cooperative effects in which the microscopic interactions have the effects of enforcing long-range order, when a material behaves as if the range of the physical force between a particle and the environment of which it is a part is literally infinite.

So what happens in, for example, a metal, which may be taken, for the sake of argument, to be a lattice of positively charged ions embedded in a sea of negative electricity? Simple calculations based on the assumption that the ions interact with each other by some defined inter-ionic potential energy will not wash, because the electron sea is of necessity a long-range system whose collective properties and, in particular, energy may not be simply expressed by the distance between the members of an arbitrarily chosen pair of atoms.

Ideally, people wishing to calculate the binding energy of a metal should proceed differently: first, make an assumption about the geometry of the ion lattice, then calculate the allowable quantum states of electrons, finally allow that the electron distribution will seek a lower level of energy in which ionic charges are more completely shielded from each other (which is where the Thomas-Fermi and other self-consistent field calculations come into their own). The results are reasonable enough, although not especially persuasive, for one thing because nobody is confident of being able to adapt the self-consistent calculations to the real case in which the ions of the ion-lattice are in motion, as they must be.

That is why there is such ample room for more empirical ways of tackling the problem by the invention of novel ways of describing the forces between ions in, say, a metallic crystal. The obvious difficulty is that the bulk properties of metals do not provide a sufficiently stringent test of possible refinements of assumed interatomic force laws, all of which must have certain general properties in common. But distortions at the surfaces of metallic crystals to the regular distribution of ions in the bulk may be an opportunity for verification, which is where Ercolessi and his colleagues begin, with the surface structure of metallic gold.

Like many other solid lattices, but more markedly, the surface of a metallic gold crystal is not a simple projection onto a crystallographic surface of the structure of the crystal lattice in bulk. Instead, the

atoms of this structure, whose unit cell is face-centred cubic, appear on simple crystallographic surfaces — such as (001) — to be arranged in stripes, five rows of atoms wide within which atoms are arranged on a triangular rather than a square pattern. This has been shown experimentally by, for example, low-angle electron diffraction.

Ercolessi *et al.* apply to this structure a force-law they have developed in other connections, one that allows not only for pair-wise forces between atoms (including other than pairs of nearest neighbours) but for a generalized glue-like force which has the effect of making it energetically advantageous, in a system such as a metallic crystal, that the coordination number should be increased. Their two-body force is conventional enough, implying repulsion at short distances, a minimum energy at about 0.27 nm and attractive forces at greater distances. The glue force is harder to visualize, but represents a kind of trade-off between high coordination number and proximity between pairs of atoms.

The outcome is remarkably suggestive even though it leaves much to be desired. By a suitable choice of the empirical constants in the force laws, five-row stripes of atoms can indeed be made to appear on the simple surfaces of gold crystals. The obvious snag is that there seems very little hope of being able to calculate either the pair-wise forces or those that represent the glue from what might be called first principles. Even so, especially because of the importance of, for example, the surface structure of semiconducting materials such as silicon, there is every prospect that this new technique of calculation will quickly become fashionable. It may even become a fashion that will last.

The situation in the calculation of the properties of metals, in other words, may soon not be very different from that in the calculation of the properties of gases half a century ago. Then it was that people were aware of the complexity of the pair-wise interatomic force as well as of the need to pay some attention to three and multi-body forces. The force-laws were for the most part empirical, but it is remarkable how useful they proved to be. Nobody will weep if there is a period, perhaps a long one, during which the calculation of the properties of metals is as much a matter of finding, through trial and error, which methods work and which do not.

John Maddox

Neurochemistry

Lessons from large molecules

from William S. Agnew

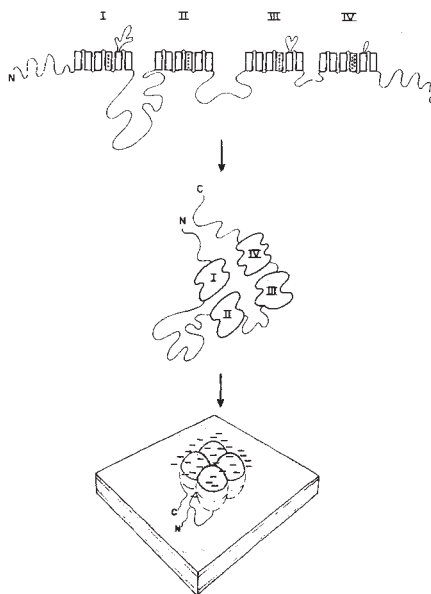
THE nervous impulse is one example of a major, easily observable physiological phenomenon that results from the action of a particular type of protein molecule, in this case the voltage-sensitive sodium channel. The nature of the smallest molecular unit that constitutes a functional sodium channel is, therefore, of considerable interest. In the most recent work from the laboratory of S. Numa, reported by M. Noda and co-workers on page 826 of this issue (ref. 1, see also refs 2,3), the authors use molecular cloning techniques to provide convincing evidence that the minimum molecular element constituting a sodium channel is a large, heavily glycosylated polypeptide of relative molecular mass 260,000–295,000 (260–295K). This finding helps to constrain models for ionic permeation and voltage-dependent gating mechanisms, and has probably resolved the question of the obligatory role of the small peptides found in preparations of brain and muscle channels.

The origin of this question lies in studies in which sodium channels have been isolated, mainly from three tissue sources: eel electroplax^{4,5}, mammalian muscle⁶ and brain⁷. Common to all three preparations are the 260–295K glycopeptides. Brain channels include two smaller glycopeptides, 36 and 33K, the former linked to the larger peptides by disulphide bridges. There is at least one peptide of 39K in muscle, although no small peptides have been found in the electroplax protein.

Each of these preparations, including the single-peptide electroplax protein, has been successfully reconstituted in biochemical and biophysical studies^{8–10}. Because of the similarities of the large peptides, these findings appear to argue against a mechanistic role for the smaller subunits. Nevertheless, evidence from subunit-dissociation studies led Catterall and co-workers to conclude that they are involved in neurotoxin binding¹¹, suggesting that they are essential in channel activity *in vivo*, perhaps for activity or for expression at the cell surface.

The three recent reports from Numa's laboratory^{1–3} cast the whole constellation of questions about sodium-channel structure and function in a new light. In the first of these² the authors cloned and sequenced the 8-kilobase complementary DNA that encodes the large peptide of electroplax. The deduced amino-acid sequence is consistent with previous compositional studies⁴ and provides an intuitively satisfying indication that the 208K polypeptide portion of the glycopeptide probably folds into a tetrameric structure.

But messenger RNA transcribed from this DNA apparently does not yield functional channels when microinjected into frog oocytes, possibly because of failure of post-translational modifications. Thus, in a second study³, Numa's group cloned the large subunit of the brain channel, only to discover that in this tissue not one, but at least three species of channel are expressed, closely related both to one another



Sodium channel amino-acid sequences suggest that a tetrameric bracelet of domains (I–IV) contain most glycosylation sites and predicted membrane-spanning α -helices. Negative surface charges derive from sialic acids and a possible cytoplasmic domain could act as a substrate for protein kinases or mediate interactions with other proteins.

and to the electroplax protein. These findings at least briefly raised the spectre that there could be as many as six smaller peptide constituents: identifying which are associated with and required for the activity of each large peptide might have demanded an entire new round of detective work before mutagenesis studies could begin.

In the study reported in this issue¹, the authors prepared messenger RNA by *in vitro* transcription of each of the three variant brain clones. When the RNA from one of these, rII, is microinjected into frog oocytes, comparatively large voltage-regulated sodium currents appear after 3–5 days of incubation. These currents are tetrodotoxin-sensitive, sodium selective and exhibit voltage activation and inactivation. Activation- and inactivation-gating parameters are within the range expected for conventional sodium channels. Expression is not enhanced by ad-

dition of low molecular mass poly(A)⁺ messenger RNA from brain, consistent with earlier fractionation studies¹². The possibility that the protein has combined with endogenous small peptides still needs to be tested directly, but seems unlikely because oocyte sodium channels are markedly different from those produced by RNA injection.

Thus, at least one of the brain-derived large peptides is sufficient to form a traditional, voltage-responsive sodium channel. In turn, because of the similarities in the sequences and the biophysical and pharmacological properties of sodium channels from different tissues, many of the structures of greatest initial interest (such as the ion pore, voltage-sensing elements, sites of interaction with neurotoxins and anaesthetics) are probably all specified within the large polypeptides. The use of single-channel recording techniques, together with genetically altered channels or with chemically modified reconstituted molecules, should help in the analysis of these essential proteins.

What, then, are some of general structural features of sodium channels that are emerging? First, within the amino-acid sequences of all of the peptides there are four very similar stretches, perhaps reflecting replication of a primordial gene (see figure). These regions (I–IV), each of about 30K, contain all the sequences predicted to form membrane-spanning α -helices. Among these are intriguing periodic sequences in which two hydrophobic residues are followed by an arginine or lysine residue for 4–8 cycles, inviting speculation about gating mechanisms^{13,14}. The intervening sequences, together with the carboxy and amino termini, are polar and not expected to span the membrane. The homologous repeats can be compared with the subunits of the pentameric acetylcholine receptor, suggesting pseudo-subunits on a string which may form a tetramerically ordered bracelet structure (see figure).

The repeat domains are very closely conserved between the electroplax and brain channels, whereas the intervening

1. Noda, M. *et al.* *Nature* **322**, 826 (1986).
2. Noda, M. *et al.* *Nature* **312**, 121 (1984).
3. Noda, M. *et al.* *Nature* **320**, 188 (1986).
4. Miller, J.A., Agnew, W.S. & Levinson, S.R. *Biochemistry* **22**, 462 (1983).
5. Norman, R.L., Schmid, A., Lombet, A., Barhanian, J. & Lazdunski, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4164 (1983).
6. Barchi, R.L. *J. Neurochem.* **40**, 1377 (1983).
7. Hartshorne, R.P., Messner, D.J., Coppersmith, J.C. & Catterall, W.A. *J. biol. Chem.* **257**, 13888 (1982).
8. Weigle, J. & Barchi, R.L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3651 (1982).
9. Talvenheimo, J.A., Tamkun, M.M. & Catterall, W.A. *J. biol. Chem.* **257**, 11868 (1982).
10. Rosenberg, R.L., Tomiko, S.A. & Agnew, W.S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1239:5594 (1984).
11. Messner, D.J. & Catterall, W.A. *J. biol. Chem.* **261**, 211 (1986).
12. Sumikawa, K., Parker, I. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7994 (1984).
13. Guy, H.R. & Seetharamulu, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 508 (1986).
14. Catterall, W.A. *Trends Neurosci.* **91**, 7 (1986).

and tail sequences are far more divergent, suggesting that the highly conserved biophysical and pharmacological properties can be accounted for largely by domains within the repeat segments. Interestingly, the brain proteins have a large (19–20K) sequence between repeats I and II, containing several consensus phosphorylation sites, that is not present in the electroplax protein. This domain could be involved in modulation, in interactions with smaller subunits or in cytoskeletal attachment. It probably does not participate directly in the 'activity cycle'. In addition, all the peptides are heavily glycosylated; in the electroplax protein sialic acids contribute up to 120 negative charges to the extracellular surface although there may be less in brain and muscle proteins. The new work suggests that for this second ion-channel

prototype the striking departure from the multi-subunit design of the acetylcholine receptor may be more apparent than real.

Also worth mentioning is the discovery of a family of sodium-channel proteins in brain. Some of these may be specific for particular neurones, and some neurones may exhibit multiple species of channel; some channels may be associated with glial cells. If there are banks of genes for ionic channels, their expression in specific cells at appropriate topographical locations must involve regulatory mechanisms that influence and perhaps even respond to the flow of electrical information in neurones and neuronal networks. □

William S. Agnew is in the Department of Physiology at Yale University School of Medicine, New Haven, Connecticut 06510, USA.

Archaean geology

Hungry komatiites and indigestible zircons

from G.R. Edwards and E.G. Nisbet

THERE has been much recent academic interest in the rocks hosting the Kambalda nickel deposit in Western Australia. This is presumably not because salaries are so low that research workers are hoping to find a new mine, but rather because these rocks are forcing a reassessment of one of the most important dating methods in old rocks and are causing us to think again about the processes that occurred when hot lavas erupted through and onto the continental crust 2,700 million years ago.

At Kambalda, nickel sulphide ores are closely associated with komatiites, lavas with high magnesium contents that generally occur in rocks older than 2.5×10^9 years. The high Mg content implies that komatiites were very hot (up to $1,650^\circ\text{C}$) when they erupted. Interest in the Kambalda rocks has concerned the age of the komatiites as well as the way in which the sulphide deposits were formed. The debate began when an array of Sm–Nd isotope data from Kambalda was interpreted as evidence that the Kambalda lavas are 3.2×10^9 years old¹. Another group interpreted the same data as the result of mixing of material derived from a depleted mantle source with either an undepleted mantle source or material from the lower continental crust². This group went on to show that Pb–Pb isotopes defined an age of 2.73×10^9 years, and that Pb isotopes in the rocks may be dominated by the effects of contamination with crustal or surface-derived material.

Compston *et al.*¹ now provide additional evidence from zircons that not only supports the 2.7×10^9 -year estimates but also

provides further evidence for crustal contamination. Zircons separated from basalts at Kambalda and analysed for U, Th and Pb isotopes using the SHRIMP ion microprobe have distinct cores with minimum ages ranging from greater than 3,400 to 3,100 million years and surrounded by mantles of younger zircon.

The old zircons are thought to be xenocrysts, partly because there were also zircon mantles and whole zircons with ages less than 2,700 million years in the sample, but also because the zircons are rich in U and Th, suggesting crystallization from felsic magma or recrystallization in felsic rock rather than crystallization from basalt. Mantling of the structural cores of the complex zircons seems to have occurred in two steps: (1) metamorphic recrystallization of the zircons while in their parent rock about 3.1×10^9 years ago, probably under granulite conditions; and (2) new growth around the metamorphically mantled zircon cores after their inclusion in the mafic–ultramafic magma. From their analysis of younger zircons, Compston *et al.* suggest the age of extrusion of the lavas is $\leq 2.669 \pm 11$ million years.

What is the ultimate source of the zircons? They are soluble in high-Mg basaltic magmas and would be unlikely to survive a long journey to the surface unless protected by enclosure in small xenoliths. It is also unlikely that new zircons or zircon mantles will grow in zirconium-undersaturated high-Mg basaltic magmas. Nevertheless, zircons are notoriously resistant to melting, so they may survive these processes virtually unscathed and

the mantles may thus represent an attempt at solid-state re-equilibration by zircons derived from the continental crust. Alternatively, some of the zircons may have been locally derived by assimilation of some of the zircon-bearing interflow sediments. The weight of the isotope evidence thus suggests that there was some contamination of the Mg-rich magmas at Kambalda. Contamination is greatest when heat transfer to the country rock is most rapid, as is the case during turbulent flow in hot, low-viscosity komatiitic magma³.

Ascending hot komatiitic magmas could have thermally eroded the wall rocks of conduits to become contaminated: on eruption, komatiites could have eroded underlying strata. Trace-element evidence⁵ implies that komatiite at Kambalda was contaminated by thermal erosion and assimilation of a mixture of sediment and tholeiites, whereas the high-Mg basalts at Kambalda are the result of up to 25 per cent contamination of komatiite, at depth, by material similar to upper continental crust. As for the nickel sulphide deposits — the reason why so many geologists descended on Kambalda in the first place — Groves *et al.*⁶ concluded from field evidence that thermal erosion occurred, but only where concentrations of sulphide liquid (which has a high thermal conductivity) were present at the base of lava channels. Some of this sulphide liquid may have formed from initial thermal erosion of sediments.

What can we learn from all this? The first lesson is that Sm–Nd data from komatiite rocks must be treated with great caution. On the other hand, zircons are extremely interesting if one has an ion microprobe. More generally, the controversy may help to elucidate the mechanisms of komatiite eruption and also the history of the continental crust, which has been sampled by the erupting liquids. Thus some zircon-containing lavas may be useful geochemical probes of the ancient lower continental crust. But some komatiites (especially the most magnesian) may be little contaminated. Perhaps they erupted through conduits that split slightly older komatiite dykes and were erupted on top of komatiite flows. Other komatiitic basalts may be the highly contaminated end-products of the eruption of originally highly magnesian liquids. And as for the nickel deposits, perhaps erosional effects under komatiites are a sign of a nearby sulphide pool. □

1. Compston, W. *et al.* *Earth planet. Sci. Lett.* **76**, 299 (1986).

2. Clauue-Long, J.C. *Nature* **307**, 697 (1985).

3. Chauvel, C. *et al.* *Earth planet. Sci. Lett.* **74**, 315 (1985).

4. Huppert, H.E. & Sparks, R.S.J. *Earth planet. Sci. Lett.* **74**, 371 (1985).

5. Arndt, N.T. & Jenner, G.A. *Chem. Geol.* (in the press).

6. Groves, D.I. *et al.* *Nature* **319**, 136 (1986).

G.R. Edwards and E.G. Nisbet are in the Department of Geological Sciences, University of Saskatchewan, Saskatoon S7N 0W0, Canada.

Antibody diversity

New mechanism revealed

from Frederick W. Alt

THE production of a functional antibody molecule requires the rearrangement of separate gene segments within the germ line of B lymphocytes. Three different segments are rearranged to produce the antigen-binding variable region of the antibody heavy chain. It was thought that no further rearrangements of completely assembled variable-region genes can occur, but two groups^{1,2} report on pages 840 and 843 of this issue that permanent cell lines of the B-cell lineage can exchange most of the body of a rearranged heavy chain variable-region segment with that of an upstream, germ-line variable region gene segment. This novel rearrangement has implications for the effi-

ciency of B-cell differentiation and establishes a new mechanism for generating the antibody diversity essential for a competent immune system.

The three germ-line DNA elements of the variable (V) region of an immunoglobulin heavy chain: variable (V_H), diversity (D) and joining (J_H) segments, are found on the same chromosome. In the mouse four J_H segments are followed immediately upstream by about 12 D segments that are in turn followed at an unknown upstream distance by 200 or more V_H segments³. Recombination between these elements is mediated by conserved recognition sequences that consist of a palindromic heptamer (related to the sequence CACTGTG) and an AT-rich nonamer separated by a spacer of either 12 or 23 base pairs (bp). Joining appears to occur only between elements flanked respectively by recognition sequences having spacers of 12 and 23 bp (the 12/23 rule)⁴. This requirement restricts the possible rearrangements as J_H segments are flanked upstream by recognition se-

quences containing 23-bp spacers. D segments are flanked on both sides by 12-bp spacers, and V_H segments are flanked downstream by recognition sequences with 23-bp spacers (see figure). In accord with the 12/23 rule, assembly of the V_HDJ_H gene occurs by joining of D and J_H segments followed by joining of a V_H segment to the pre-existing DJ_H complex⁵. A consequence of the 12/23 rule is that a pre-formed DJ_H rearrangement can be replaced by joining an upstream D to a downstream J_H but a V_HDJ_H rearrangement, once formed, cannot be replaced by joining events to downstream J_H segments because direct V_H to J_H joining is prohibited by the 12/23 rule and all intervening D

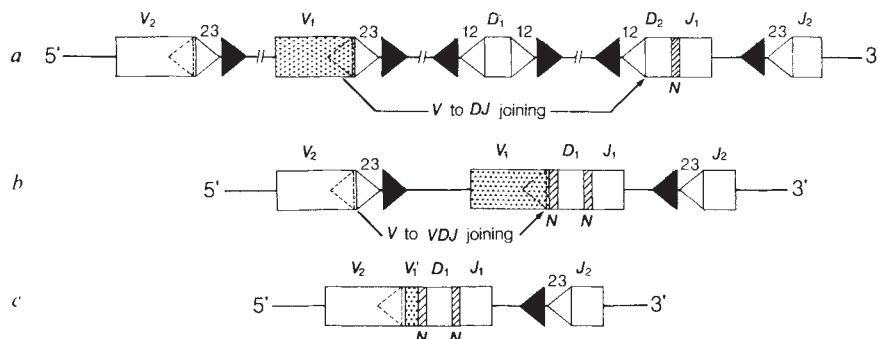
segments (12-bp recognition sequences are deleted during V_HDJ_H assembly)³. The novel V_H to V_HDJ_H rearrangement described by Reth *et al.*¹ and by Kleinfeld *et al.*² in this issue occurs by a recombination mechanism similar to that used for normal V-gene assembly but uses a recognition heptamer (TACTGTG) found in the 3'-region of most V_H genes⁶ (see figure). Notably, this heptamer is identical to that within the recognition sequence flanking the 5' border of most D segments so that although the heptamer 5' of the D segment is deleted after V_H to DJ_H joining, an identical heptamer within the V_H gene is now present. The V_H gene sequence does not contain an obvious nonamer so the complete upstream recognition sequence is not replaced. However the structure of the V_H to V_HDJ_H recombinants suggests that the internal heptamer mediates site-specific joining, rather than conversion or homologous recombination events.

Site-specific immunoglobulin-gene recombination mediated by an isolated heptamer has been shown previously in the light chain^{7,8}. In addition, although the TAC and TGT codons of the heptamer encode a conserved Tyr and Cys found in most V-gene segments, Kleinfeld *et al.* note little third-base variation within these residues in V_H gene segments, which they believe suggests conservation of the sequence itself. This is perhaps surprising in the light of recent evidence that immunoglobulin recombinase accepts wide sequence variation within the heptamer, at least when it is associated with a nonamer¹⁰. But Reth *et al.* indicate that V_H to V_HDJ_H joining occurs frequently even though direct V_H to J_H joining does not occur, perhaps suggesting that the nonamer restricts recombination in the context of the 12/23 rule. Clearly more remains to be learned about the role of recognition sequences.

Assembly of the elements in the V-region gene is not precise and bases can be lost and/or added at the joining point (N regions)³. The V_HDJ_H joint region of the antibody molecule is involved in antigen contact; thus, junctional diversification provides antibody diversity. The studies indicate that V_H to V_HDJ_H joining may allow additional diversification of this important region: sequences left attached from the first V-region provide an additional N-like region (see figure).

Junctional diversification mechanisms have another important consequence for differentiating B lymphocytes. The reading frame of the heavy-chain messenger RNA is determined by the start site for translation at the 5' end of the V_H . Because of random insertion and deletion of bases and the triplet reading frame, approximately 2 out of 3 V_HDJ_H joins link the V in a different translation reading frame from the J. Such rearrangements are 'aberrant': they produce heavy-chain mRNA that cannot be translated into a complete heavy-chain protein. Given the inaccuracy of the joining process, in spite of the opportunity for rearrangement on both chromosomes, nearly 50 per cent of differentiating B cells should join both heavy-chain alleles out of frame³. Previously, such cells were thought to be lost from the differentiation pathway, because aberrant V_HDJ_H joints could not be salvaged. However, Reth *et al.* demonstrate that permanent pre-B lines that have formed two aberrant V_HDJ_H rearrangements can rescue the heavy-chain production by V_H to V_HDJ_H recombination. Because of the same imprecise recombination this secondary join can provide a productive reading frame and could make this phase of B-cell differentiation nearly 100 per cent efficient.

The production of complete heavy chains is restricted to a single allele in B cells. Such allelic exclusion appears to result from cessation of V_H to DJ_H rearrangement when the initial V_H to DJ_H join



a. Scheme for the chromosome after a DJ_H join. Recognition heptamers are indicated by open triangles and nonamers by closed triangles; 12- and 23-bp spacers are indicated. The heptamer within the V segment is represented by a dashed triangle. b. A V_H to DJ_H join; intervening D segments (and 12-bp spacers) are deleted; added N segments are indicated. c. A V_H to V_HDJ_H join; in this case the final assembled V gene consists of the body of V_H2 , a small segment of V_H1 , a 5' N region, a D segment, a 3' N region, and a J_H segment. Sizes of the segments not drawn to scale.

ciency of B-cell differentiation and establishes a new mechanism for generating the antibody diversity essential for a competent immune system.

The three germ-line DNA elements of the variable (V) region of an immunoglobulin heavy chain: variable (V_H), diversity (D) and joining (J_H) segments, are found on the same chromosome. In the mouse four J_H segments are followed immediately upstream by about 12 D segments that are in turn followed at an unknown upstream distance by 200 or more V_H segments³. Recombination between these elements is mediated by conserved recognition sequences that consist of a palindromic heptamer (related to the sequence CACTGTG) and an AT-rich nonamer separated by a spacer of either 12 or 23 base pairs (bp). Joining appears to occur only between elements flanked respectively by recognition sequences having spacers of 12 and 23 bp (the 12/23 rule)⁴. This requirement restricts the possible rearrangements as J_H segments are flanked upstream by recognition se-

segments (12-bp recognition sequences are deleted during V_HDJ_H assembly)³.

The novel V_H to V_HDJ_H rearrangement described by Reth *et al.*¹ and by Kleinfeld *et al.*² in this issue occurs by a recombination mechanism similar to that used for normal V-gene assembly but uses a recognition heptamer (TACTGTG) found in the 3'-region of most V_H genes⁶ (see figure). Notably, this heptamer is identical to that within the recognition sequence flanking the 5' border of most D segments so that although the heptamer 5' of the D segment is deleted after V_H to DJ_H joining, an identical heptamer within the V_H gene is now present. The V_H gene sequence does not contain an obvious nonamer so the complete upstream recognition sequence is not replaced. However the structure of the V_H to V_HDJ_H recombinants suggests that the internal heptamer mediates site-specific joining, rather than conversion or homologous recombination events.

Site-specific immunoglobulin-gene recombination mediated by an isolated

is productive. In the event studied by Kleinfield *et al.*², the V_H of a productive V_HDJ_H rearrangement is exchanged with an upstream V_H segment to produce a novel productive rearrangement, indicating that the recombinase is still active in cells with a productive V_HDJ_H . This finding may, at first glance, seem inconsistent with an allelic exclusion mechanism regulating V_H to DJ_H rearrangement. However, current evidence suggests that joining is prevented in cells with a productive V_HDJ_H rearrangement by limiting access of the recombinase to V_H gene segments rather than by turning off recombinase activity³. Recombinase is thought to lose access because transcription of germ-line V_H segments stops in cells that produce complete heavy chains⁴.

Notably, the described V_H to V_HDJ_H recombinations seem to involve proximal V_H segments and it has been shown that the activity of the heavy-chain enhancer in the J_H to D_H intron allows transcription of the proximal unrearranged V_H segment⁵; perhaps by providing access to recombinase. Because allelic exclusion appears to occur in nearly 100 per cent of normal B cells, it seems that V_H to V_HDJ_H joining does not frequently activate expression of both heavy-chain alleles.

V_H to V_HDJ_H joining may also prevent bias in the development of the heavy-chain V -region repertoire. Differentiating B cells use the V_H segments nearest the J_H elements at high frequency to form V_HDJ_H rearrangements⁶, which limits the spontaneously generated heavy-chain repertoire. But the repertoire of mature peripheral B cells appears to involve all V segments equally⁷. How this occurs is not known, although cellular selection mechanisms have been implicated. Kleinfield *et al.* suggest that V_H to V_HDJ_H joining may be a means of successively replacing rearranged 3' V segments with more 5' V segments. However, if replacement involves nearest neighbours, as suggested from the data of Reth *et al.*, it is difficult to imagine successive recombination events between hundreds of V genes playing a major role in the process. Final understanding awaits further elucidation of the mechanism and frequency of this process. □

1. Reth, M., Gehrmann, P., Petrac, E. & Wiese, P. *Nature* **322**, 840 (1986).
2. Kleinfield, R. *et al.* *Nature* **322**, 843 (1986).
3. Alt, F.W. *et al.* *Immun. Rev.* **89**, 5 (1986).
4. Early, P. *et al.* *Cell* **19**, 981 (1980).
5. Sakano, H. *et al.* *Nature* **286**, 676 (1980).
6. Early, P. *et al.* *Molec. cell. Biol.* **2**, 829 (1982).
7. Seidman, J. & Leder, P. *Nature* **286**, 779 (1980).
8. Durdik, J., Moore, M.N. & Selsing, E. *Nature* **307**, 749 (1984).
9. Siminovitch, K.A. *et al.* *Nature* **316**, 260 (1985).
10. Yancopoulos, G.D. *et al.* *Cell* **44**, 251 (1986).
11. Yancopoulos, G.D. & Alt, F.W. *Cell* **40**, 271 (1985).
12. Wang, X.-F. & Calame, K. *Cell* **43**, 659 (1985).
13. Dildrop, R. *et al.* *Eur. J. Immun.* **15**, 1154 (1985).

Frederick W. Alt is in the Department of Biochemistry and Molecular Biophysics at the College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA.

Animal behaviour

How great white sharks, sabre-toothed cats and soldiers kill

from Jared M. Diamond

PREDATORS that attack large prey endowed with formidable defensive weapons face a dilemma. If the victim is not incapacitated by the first blow and the predator must continue to engage the victim in order to kill it, the predator itself risks injury. For example, lions are occasionally crippled or killed by zebra, giraffe or rhinoceros prey. Recent studies suggest that two quite different large predators — the great white shark^{1,2} and the extinct sabre-toothed cats³ — independently evolved the same solution to this dilemma.

Adult great white sharks (*Carcharodon carcharias*), the largest living marine fish capable of preying on vertebrates, often feed on seals, whales and other sharks. Common victims in Californian waters are elephant seals, which have formidable teeth combined with a maximum size exceeding that of the shark. The seal is more agile than the shark and is therefore most efficiently hunted by surprise. Timothy Tricas and John McCosker^{1,2} reconstructed the shark's strategy from observations of attacks; scars on seals that got away; and underwater observations (while protected within steel cages) of sharks approaching bait. Bite scars on surviving seals indicate a single massive bite on the underside of the body. Apparently a shark cruising underwater sees above it the silhouette of a seal basking at the surface, rises undetected towards the seal from behind and below, quickly bites out a large chunk of the seal (the entire biting takes only 1 second), and rapidly retreats to avoid injury. The shark then waits for the seal to go into shock or to haemorrhage before closing in to complete the kill and to feed.

This reconstructed sequence helps one to understand what happens to human victims of great white sharks. Most survivors never saw the shark that attacked them: they merely

felt themselves lifted into the air and then dropped after suffering a single haemorrhaging bite. Victims usually die from loss of blood, not of limbs or vital organs.

A similar strategy appears to have been practised on land by extinct sabre-toothed cats of the genera *Smilodon* and *Homotherium*³. These lion-sized predators take their name from their very long upper canines (exposed tooth up to 15 cm long in *Smilodon*). How these weapons were used to kill has been in dispute. The prevalent view until recently was that the upper canines were used to stab or slash the victim while the lower jaw was merely drawn down out of the way. However, closer examination casts doubt on the usefulness of the blunt sabres for stabbing and slashing: they are round-tipped, oval rather than knife-like in cross-section, and barely clear the lower jaw.

Our most detailed information about the preferred prey of sabre-tooths comes from a probable den, Friesenhahn Cave in Texas, where remains of adult plus very small juvenile *Homotherium* co-occur with scores of baby mammoths, presumably killed by the adult cats and dragged to the den to feed the kittens. Even a baby mammoth could hardly have been killed by the cat at a single blow. Yet it would

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

How *Smilodon* used its peculiar anatomy to rip a large piece of flesh out of its prey. Propping its mandible on the prey (a), the cat used its powerful head-depressing muscles to drive in the upper canines (b), rotate the head until the jaws closed (c, d) and pull out a fold of flesh. (Drawing by Mark Hallett. From ref. 3, courtesy of the Los Angeles County Museum of Natural History.)

have been suicidal for the cat to continue to grip the baby, as the mother would then have time to crush or impale the cat.

A solution to this dilemma is suggested by William Akersten's anatomical studies³ of remains of *Smilodon* trapped in the famous La Brea tar pits of Los Angeles. Muscle scars on *Smilodon* bones show that the muscles for closing the jaws were not especially strong but that the muscles to depress the head by rotating it downwards around the joint with the first neck vertebra were massive (see figures). The chin was flanged, and the jaws opened to the enormous gape angle of 95° to permit the same clearance between the long upper and lower canines as does the 65° gape of modern cats. Akersten suggests the following sequence. A cat ambushes a baby mammoth straying from its mother and uses its powerful fore-limbs to pull down the baby, exposing the abdomen. The cat opens its jaws wide, catches a large fold of flesh between the upper and lower



(From ref. 3.)

Cranium and mandible of *Smilodon*.

canines, wedges the lower jaw firmly against the prey by the chin flanges, and then completes the bite by the head-depressing muscles. It then flees from the approaching mother mammoth and waits for the baby to haemorrhage, die, and be abandoned by its mother.

Yet one more predator has independently evolved a similar solution to the problem posed by a formidable victim. The most dangerous prey of all is an armed man, routinely hunted only by other armed men. Inexperienced soldiers aim for the head or heart, but these small targets are easily missed. "Always aim for the stomach when you shoot" was the advice of an experienced African bush fighter — "it is just as good, because no man can live long after his intestines have been shot away". Nor could an elephant seal or baby mammoth. □

1. Triacas, T.C. & McCosker, J.E. *Proc. Calif. Acad. Sci.* **43**, 221 (1984).

2. McCosker, J.E. *Calif. Acad. Sci. Mem.* no. 9, 123 (1985).

3. Akersten, W.A. *Los Ang. County Mus. nat. Hist. Contr. Sci.* no. 356 (1985).

4. Bosman, H.C. in *The Collected Works of Herman Charles Bosman* 64 (Bali, Johannesburg, 1981).

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Anthropology

The longest human record

from John E. Yellen

THE recent conference* to honour J. Desmond Clark for his contributions to African pre-history provided a unique opportunity to review current ideas about the archaeological record of this continent. African palaeoanthropology is most clearly distinguished by its quest for the first uniquely hominid antecedent: the earliest known hominid fossils derive from the Eastern Rift Valley in Kenya. Evidence for the first million years of cultural development, in the form of stone tools, butchered animal remains and possibly the controlled use of fire, are also limited to the African continent. The first modern humans, *Homo sapiens sapiens*, were widely distributed in Africa well before their appearance in Europe and the near East and the development of complex societies, in relative independence from those in other parts of the Old World.

Excavations this past summer near Ishango, eastern Zaire, produced stone artefacts associated with faunal remains that are just over two million years old (Jack Harris, University of Wisconsin; Noel Boaz, Virginia Museum of Natural History). In conjunction with stone-tool assemblages from Ethiopia of about the same age, systematic lithic modification must have appeared more than two million years ago. Unfortunately, these artefacts cannot be assigned to specific species.

Arguments were presented that *Homo habilis* was capable of speech (Dean Falk, University of Puerto Rico; Philip Tobias, University of the Witwatersrand). This species had a larger brain than australopithecine forms with prominent enlargement of Broca's and Wernicke's areas, both associated with language in modern humans. These advances appeared at about the same time as the first stone tools.

Whereas most palaeoanthropologists agree that by the mid-Pleistocene *Homo erectus* had controlled use of fire, there has been no evidence for earlier occurrences. Because charcoal decomposes in tropical environments, African evidence is extremely difficult to obtain. Recently discovered patches of reddened earth in East African sites dates between 1.5 and 1.7 million years ago and which are associated with cultural remains were suggested to indicate much earlier human use of fire (Desmond Clark, University of California; Jack Harris, University of Wisconsin). But palaeomagnetic and thermoluminescent analyses failed to provide unequivocal evidence of sediment heating.

* The Longest Record: The Human Career in Africa. Berkeley, California. 12 – 16 April 1986.

The question of early hominid hunting reflects on cognitive ability and social organization. Although animal bones with stone-tool cutmarks provide evidence of carcass use by humans earlier than 1.5 million years ago, it is uncertain whether these remains were acquired by hunting or by scavenging from carnivore kills. If hominids were primarily scavengers, they would be left with the less-choice body parts, which would be reflected in the skeletal elements recovered and the way they were treated to remove remaining scraps of meat. Using such criteria, faunal remains from both Olduvai and Koobi Fora can be interpreted to indicate hunting as cutmarks appear on the primary meat-bearing bones of small bovids which would be unlikely to survive the attention of large carnivores (Henry Bunn, University of Wisconsin).

Also at issue was a large sample of excavated cattle and sheep bones from the Neolithic site of Ngamuriak in southern Kenya. In this butchered assemblage the larger and smaller species are processed differently and a similar pattern is evident in early hominid material (Fiona Marshall, University of California). On the other hand, analysis of the Olduvai Bed I fauna shows 13 cases of overlapping stone-tool and carnivore tooth marks. In eight of these the tool mark overrides that made by tooth, indicating human scavenging (Pat Shipman, Johns Hopkins University).

Recent work on the eastern margin of the Kalahari Desert shows a pattern of increasing complexity and hierarchical organization which, in effect, set the stage for the emergence of the Zimbabwe culture (James Denbow, National Museum of Botswana). The Venda, who are among the lineal descendants of Zimbabwe people, use variation in motifs on compound walls to denote status. Direct counterparts exist in Zimbabwe period stone-walled structures and, by application of Venda data, it has been possible to reconstruct a five-level hierarchy which extended from heads of individual households to the divine national leader (Thomas Huffman, University of the Witwatersrand). Because the gradual evolution of this system can be traced through time it is highly unlikely that the final structure was a direct results of outside stimulus (trade with Indian Ocean powers) but more probably derived from indigenous roots. □

John E. Yellen is in the Anthropology Program at the National Science Foundation, 1800 G Street NW, Washington DC 20550, USA.

Neurophysiology

Parallel channels and redundant mechanisms in visual cortex

from N. V. Swindale

ALL too rarely, pharmacologists come up with a drug that interferes selectively with a particular pathway or a single well-defined function of the nervous system. Ideally the drug should be easily administered, specific and reversible, but such combinations are not often achieved. One molecule that does seem to satisfy these criteria is the glutamic acid analogue 2-amino-4-phosphonobutyric acid (APB) which has been studied by P.H. Schiller and collaborators and whose latest work is reported on page 824 of this issue¹. APB blocks the function of a class of retinal ganglion cells, and it is hoped that it will help to determine to what extent the visual system functions as an integrated whole or whether its different components are capable of some degree of independent function.

The pathway from retina to brain is divided into parallel channels, each selective for different attributes of the image, such as light intensity, colour and spatial frequency. Each channel consists of neurones of two types, ON and OFF, the former responding only to light levels that are greater than the local average in the image and the latter responding (by an increase in firing rate) to levels that are less than the local average. Thus, ON channels carry information about bright spots or areas in the image and OFF channels do the reverse. This division of labour, which resembles that of the push-pull output stage of an amplifier probably has more than one advantage: it allows cells to have low levels of activity when the retina is in darkness or uniformly illuminated, reducing metabolic requirements, and it effectively doubles the dynamic range over which neurones can signal information about light levels in the retinal image.

The separation between ON and OFF pathways occurs at the synapse between the rod and cone photoreceptors and bipolar cells, the earliest stage in the visual pathway. All photoreceptor cells hyperpolarize in response to light, but the bipolar cells to which they are connected are of two types: ON bipolars that depolarize when the photoreceptor to which they are connected hyperpolarizes, and OFF bipolars that hyperpolarize when the photoreceptor hyperpolarizes. If APB is applied to the retina it hyperpolarizes the ON bipolars, silencing cells post-synaptic to them². The action of APB appears to be the same in retinas from all species of vertebrate, and is specific to all classes of

ON channel.

One of the first observations made with the drug³ was that it did not abolish the antagonistic centre-surround organization of either retinal cells or lateral geniculate nucleus cells to which the retina projects. This means that the inhibitory actions involved in generating the surround must occur entirely within the OFF pathway. The same is probably true for the ON pathway. ON and OFF channels thus remain separate in the retina and the lateral geniculate nucleus, and the pathways do not converge until the inputs reach the visual cortex.

The ability to disable half of the visual system in this way raises obvious questions. Can each half function on its own and, if so, how well? Because an edge normally has both light and dark components, it might be thought that the detection of edges and their orientation (a function of the visual cortex) will depend on the proper functioning of both ON and OFF channels. Blocking the ON channels by applying APB to the retina might be ex-

pected to interfere seriously with vision. But it appears that, although APB does have some deleterious effects, they are much less than had been supposed.

The carefully controlled experiments on awake monkeys reported in this issue show that, when treated with APB, the animal cannot detect a light coming on, although it can still detect the onset of a dark spot. Perhaps more relevant is that many visual functions are relatively unimpaired by APB. For example, the report¹ shows that monkeys treated with the drug have qualitatively normal stereopsis when tested with random-dot stereograms and can detect apparent motion in random-dot kinematograms. Both these tasks are normally regarded as rather severe tests of visual performance. Other studies³ show that cortical neurones in APB-treated animals, although lacking in ON responses, still possess orientation and direction selectivity.

Although interactions between ON and OFF channels are not necessary for the generation of many neuronal-response properties, they do not prove that such interactions do not occur, or that they have no important functional role. A really careful comparison of the properties of cortical cells between normal and APB treated animals has yet to be carried out, and strong interactions between ON and OFF inputs to the cortex are known to exist (see, for example, ref.4). It seems

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

New discoveries of some remarkable Lower Carboniferous fish¹⁻⁵ show that they had vast toothed spines apparently suspended directly above their heads. *Stethacanthus*, a 1-m-long shark, now known from nearly complete specimens from Scotland and Montana¹⁻³, had a remarkable spine shaped like a shaving brush just behind its head. The spine stands nearly vertical and is topped by dozens of small teeth. It is clearly an unwieldy structure and must have had an important function. The spine was present in both males and females, and Zangerl³ suggests that it was used to scare off potential predators: *Stethacanthus* had a patch of teeth on top of its head as well and, when partially buried in the mud, the two toothed areas would have mimicked a vast gaping mouth, as if of some giant predator. The figure illustrates the possible appearance of the fish³.

A second stethacanthid, *Falcatus*⁴, is a small shark up to 14.5 cm long and has a long, shelf-like spine extending from roots deep in the muscles of the 'shoulder' region to run over the head, like a sunshade. The spine is in two parts, the fixed base and the horizontal portion which slots on to it, and is present only in sexually mature males. Lund⁴ suggests that male *Falcatus* sharks aggregated before the breeding season for display-courtship rituals. The third complete Stethacanthid shark, *Damocles*⁵, is also small (up to 20 cm long) and differs from *Falcatus* in having a simpler spine with many sharp teeth.

Michael J. Benton

1. Wood, S. *Nature* **297**, 574 (1982).

2. Lund, R. *Géobios* **17**, 281 (1984).

3. Zangerl, R. *J. vertebr. Paleontol.* **4**, 372 (1984).

4. Lund, R. *J. vertebr. Paleontol.* **5**, 1 (1985).

5. Lund, R. *J. vertebr. Paleontol.* **6**, 12 (1986).

Michael J. Benton is in the Department of Geology, The Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland.

very likely that in many cases these interactions must be involved in generating orientation selectivity (for example, in cells that are believed to receive inputs from only a single ON centre and a single OFF centre cell, side by side).

Other recent results support the idea that orientation selectivity is produced by more than one mechanism, and in more than one location in the visual cortex. Inactivation of the A layers of the lateral geniculate nucleus (a major source of visual input to the cortex) abolishes all responses in the middle layers of the cortex, but leaves orientation selectivity in the upper cortical layers unimpaired⁵. Conversely, inactivating the upper cortical layers by cooling leaves orientation selectivity in the middle layers of the cortex intact⁶. It had previously been thought that orientation selectivity in the upper layers was merely a passive reflection of the orientation selectivity of their inputs from the middle layers of the cortex but these results^{5,6} mean this cannot be entirely true.

The idea that many cortical properties are generated redundantly, by interactions both within and between different parallel afferent pathways to the cortex, in

different cortical layers and perhaps by different mechanisms, has several attractions. Redundancy of mechanism would render the brain less sensitive to pathological or experimental damage to particular subdivisions in the nervous system. Inadequacies in any one mechanism might be compensated for by some corresponding advantage in another. More than one type of processing algorithm is needed to account for other aspects of visual performance, such as stereopsis⁷, and perhaps it is easier to evolve several relatively crude neural mechanisms that work in parallel to achieve a particular goal than to evolve just one especially effective one. Single elegant solutions to particular problems may not exist in the brain: quantity, rather than quality, of mechanism is what counts. □

- Schiller, P.H., Sandell, J.H. & Maunsell, J.H.R. *Nature* **322**, 824 (1986).
- Slaughter, M.M. & Miller, R.F. *Science* **211**, 183 (1981).
- Schiller, P.H. *Nature* **297**, 580 (1982).
- Hammond, P. & Mackay, D.M. *Expl. Brain Res.* **49**, 453 (1983).
- Malpeli, J.J. *Neurophysiol.* **49**, 595 (1983).
- Baker, F.H. & Malpeli, J. *Expl. Brain Res.* **29**, 433 (1977).
- Mitchison, G. *Nature* **304**, 123 (1983).

N.V. Swindale is in the Department of Psychology at Dalhousie University, Halifax B3H 4J1, Canada.

Chiral chemistry

Asymmetric epoxidation gives organic chemists a hand

from Martin Bryce

THE asymmetric epoxidation of allylic alcohols is one of the most innovative reactions introduced into organic chemistry in modern times¹. The tremendous potential of this reaction was recognized soon after its discovery six years ago by Sharpless and Katsuki². However, as the target molecules for organic chemists become ever more complex, several years' work may be needed to complete intricate, multistep syntheses, and it is only within the context of such challenging synthetic chemistry that the true value of a new reaction can be seen. Recently, Sharpless and colleagues have made an important improvement in the reaction conditions³, and it is now clear that the Sharpless epoxidation has come of age as a practical method for the introduction of chirality into a wide range of organic molecules.

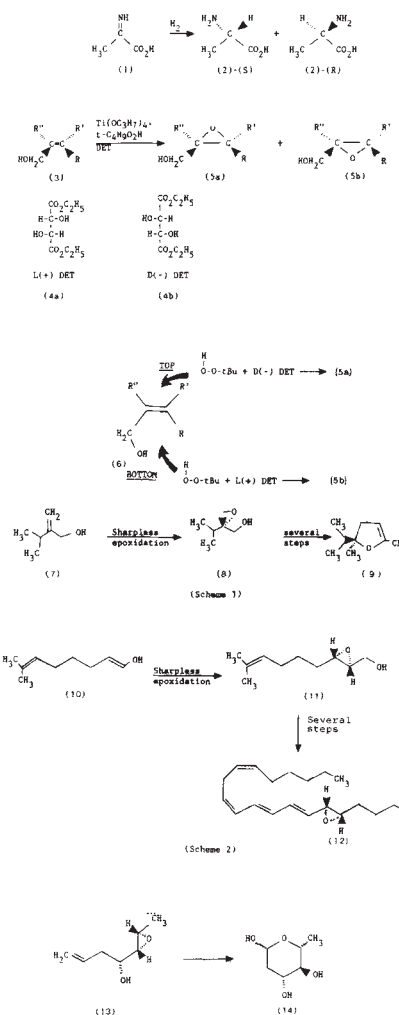
To understand the significance of this reaction we must be aware that the organic molecules of life are chiral (that is, like right and left hands, the molecules are mirror images of each other) but they are not identical because they are non-superimposable. From the most primitive to the most highly developed organisms, the life

processes at the molecular level depend critically on this chirality. A molecule which on further substitution can become chiral is termed prochiral. A chiral reagent will distinguish between the faces of a symmetrical, prochiral structure giving rise to unequal amounts of chiral products, whereas an achiral reagent will show no selectivity. For example, reduction of a prochiral amino-acid precursor (see figure: structure 1) by an achiral reagent such as hydrogen can occur with equal probability from each of the two faces of the C=N bond, giving rise to a 50:50 mixture of the two chiral forms of the amino acid (structure 2); there will be no selectivity. These chiral products are termed (S) and (R) enantiomeric forms. However, the two faces of prochiral molecule (structure 1) are not the same to a chiral reagent, such as an enzyme, which will attack one side to the virtual exclusion of the other. Indeed, for the most part, enzymes produce only the amino acids having the (2)-(S) configuration (see figure).

Organic chemists strive to develop highly efficient methods of synthesizing pure enantiomers of organic compounds,

both for academic interest and for commercial application, for example in the fields of pharmaceuticals and pesticides where, frequently only one enantiomer of a compound is biologically active. Thus, asymmetric hydrogenation using chiral catalysts has become a well-established laboratory procedure over the past decade⁴. Until recently, however, the counterpart of this reaction, asymmetric oxidation, has met with little success.

The answer has proved to be the Sharpless reaction, which is based on both organic and inorganic stereochemistry



and uses reagents that are either commercially available or easy to prepare. The reaction, which involves treatment of an allylic alcohol (see figure: structure 3) with a mixture of titanium tetra-isopropoxide, t-butyl hydroperoxide and diethyltartarate (DET) (structure 4) to give an epoxy alcohol (5), can be controlled, with (5a) or (5b) forming in more than 95 per cent enantiomeric excess. Thus, remarkably, the reaction approaches the stereoselectivity of an enzyme system and even surpasses most enzymes in the variety of substrates (groups R, R' and R'' in structure 3) that will react. In a recent improvement³, the addition of molecular

sieves to the reaction mixture enables as little as 5 moles per cent of titanium tartrate complex to be used to ensure high enantioselectivity of the epoxidation.

Sharpless considers that the titanium metal catalyst serves to organize the epoxidizing agent (t-butyl hydroperoxide), the chiral-inducing agent (diethyl tartrate) and the substrate (the allylic alcohol) into such geometry that reaction at one face of the double bond is greatly preferable to reaction at the other face¹. In the key step, the epoxide oxygen is always delivered to the same face of the olefin for reaction involving a specific tartrate isomer (structure 4a or 4b). If the allylic alcohol (structure 3) is represented by the stereochemical model (6), the oxygen is delivered from the bottom face using (+) DET and the top face using (-) DET, to give chiral epoxy alcohols (5a) and (5b), respectively.

The ability of epoxy alcohols to retain chirality while undergoing a wide range of subsequent transformations has enabled the Sharpless epoxidation to be incorporated into numerous multistep syntheses of chiral compounds (see ref. 1 for a recent review). For example, Mori's synthesis of structure 9, the sex pheromone of the beetle *Hylecoetus dermestoides* (a pest of European hardwood and softwood trees) starts from prochiral allylic alcohol (structure 7). This alcohol was transformed by the Sharpless reaction into epoxy alcohol (8) which contains the required stereochemistry for elaboration into the desired chiral product (9) (see figure; scheme 1).

A class of pharmaceuticals where improved synthetic routes incorporating the Sharpless oxidation have been described are leukotrienes, important in various inflammatory diseases. A new synthesis of leukotriene A (structure 12) involves, as a key step, formation of chiral epoxide (11) from prochiral allylic alcohol (10) (see figure; scheme 2).

The Sharpless epoxidation procedure also provides easy access to rare monosaccharides which are important structural components of certain natural products and are also valuable synthetic intermediates. For example, epoxy alcohol (structure 13) can be converted into hexose derivative (14) which is a component of the antibiotic olivomycin A.

These examples show how important asymmetric epoxidation reactions can be for those who wish to make chiral molecules with high enantiomeric purity by simple and reliable procedures. □

1. Pleninger, A., *Synthesis* 89 (1986).

2. Katsuki, T. & Sharpless, K.B. *J. Am. Chem. Soc.* **102**, 5974 (1980).

3. Sharpless, K.B. & Hanson R.M. *J. Org. Chem.* **51**, 1922 (1986).

4. Mosher, H.S. & Morrison, J.D. *Science* **221**, 1013 (1983).

5. Sharpless, K.B., Woodward, S.S. & Finn, M.G. *Pure Appl. Chem.* **55**, 1823 (1983).

Martin Bryce is a Lecturer in Organic Chemistry at Durham University, Durham DH1 3LE, UK.

Cytoskeleton

New views of the red cell network

from Velia M. Fowler

THE elaborate schemes of the molecular anatomy of the erythrocyte membrane skeleton that have evolved over the years¹⁻⁴ are based almost exclusively on studies of the properties and associations of purified proteins *in vitro*. Attempts to analyse the structure of the membrane skeleton *in situ* have been impeded by the enormous density of the components, particularly spectrin, in the plane of the membrane. Several groups^{5,6}, most recently Steck and his colleagues⁷, have succeeded in artificially spreading out the spectrin network without (apparently) disrupting the linkages holding it together. Far from being "... an anastomosing framework woven by a myopic fisherman"⁸, the membrane skeleton is organized as a strikingly regular lattice in the plane of the membrane.

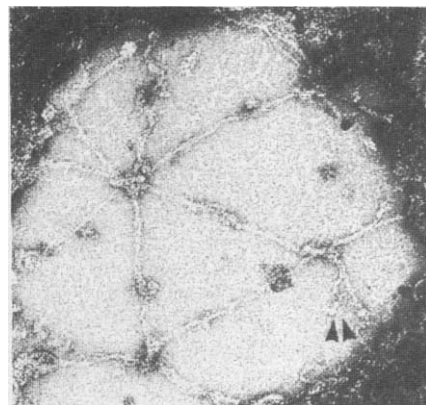
Five to seven long, thin filaments (about 180 nm long), whose morphology corresponds to purified spectrin tetramers, are attached by their ends to junctions comprised of short, stubby actin filaments (33–37 nm long), thus generating a series of polygons, mostly hexagons, in which the sides are spectrin molecules and the vertices actin filaments (see figure)^{6,7}. The latter appear thicker than normal actin filaments, implying that additional proteins are associated with them (see below). There is even a glob near the middle of the spectrin tetramers that is in the right place to be the ankyrin-band 3 membrane attachment site for spectrin^{6,7}. What is so satisfying about these negatively stained images of the spread membrane skeleton is that they are completely consistent with the models inferred from work with purified proteins.

The surface area of the spread membrane skeleton is calculated to be about 9–10 times that of the unspread membrane, indicating that the fully extended spectrin molecules (about 180 nm) are around three times longer than necessary to span the distance between adjacent actin filaments in the unperturbed membrane (about 60 nm)⁶. It is an interesting question whether the spectrin tetramers in the unperturbed membrane are actually folded or coiled into a conformation with a more compact structure, or whether they are simply looped passively from one actin filament to the next so as to accommodate the surface area of the lipid bilayer itself. This distinction could have implications for the molecular mechanism of erythrocyte shape transformations and for the viscoelastic properties of the erythrocyte membrane^{9,10}.

One source of support for the second

model is the observation that images of the spread membrane skeleton of reticulocytes are practically indistinguishable from those of the mature cell with respect to the essentially hexagonal pattern of the lattice (S.-C. Liu and J. Palek, personal communication), although the surface area is about three times greater in the one than in the other.

Thus, it seems that the reduction in membrane surface that accompanies maturation takes place without major changes in the organization of the membrane skeleton. Because analogues to spectrin, ankyrin and protein 4.1 are present on the plasma membranes of non-erythroid cells¹¹, it will be interesting to



Negatively stained preparation of the spread erythrocyte membrane skeleton (from ref. 4).

see if the organization of the membrane skeletons in these cells uses similar principles to achieve local changes in membrane area during some cell movements.

Complexes that evidently represent isolated lattice vertices and that consist of about six spectrin 'legs' extending from a body containing a short actin filament have recently been isolated¹². They also contain protein 4.1, which promotes complex formation, and protein 4.9, which is thought to be bound to actin (see refs 2–4). Tropomyosin may also be associated with the junctional complexes *in situ*, as this protein is present on the membrane in sufficient quantity to satisfy the binding capacity of the actin¹³.

A dimeric calmodulin-binding protein of relative molecular mass (M_r) 97,000/103,000 has just been discovered in the membrane skeleton. It is also thought to be associated with the junctional complexes, as there are about the right number (30,000 copies) per cell, that is 1 dimer per filament of 16 actin monomers¹⁴. But it is not yet known with which of the proteins in the membrane skeleton it is associated. Immunolocalization of these various pro-

teins in the membrane skeleton will be necessary to ascertain whether they are indeed associated with the junctional complexes that form the vertices of the spectrin lattice.

Why should the molecular substructure of the junctions in the erythrocyte membrane skeleton be so complex? An attachment site for about six spectrins would need only a small actin filament containing six actin monomers, along with six molecules of protein 4.1. However, a junctional complex composed only of spectrin, protein 4.1 and actin would not be self-limiting because of the propensity of actin to polymerize into long filaments under physiological conditions.

Restriction of actin-filament length by the accessory proteins mentioned above (alone or in combination) could serve to regulate the number of spectrin molecules attached to the junctions and thus specify the surprisingly regular lattice structure of the network. Some slippage in filament length regulation could easily account for variation in numbers of spectrin molecules attached to the junctions^{5,7}.

An apparent paradox is that although the stoichiometry of spectrin-protein 4.1 binding to actin filaments *in vitro* is equimolar, there are only about six spectrin molecules attached to any one of the 13- to 15-monomer-long actin filaments^{5,7}. Because tropomyosin binds in the two grooves of the actin-filament helix without restricting filament length¹³, and protein 4.9 bundles actin filaments *in vitro*¹⁵, we still have no explanation for the actin-filament length restriction that is observed *in situ*.

In addition to ensuring the correct assembly of the membrane skeleton, it is conceivable that changes in the structure of the junctional complexes mediate the calcium- and ATP-dependent discocyte

-echinocyte shape transformations of erythrocytes. A calcium- and calmodulin-dependent protein kinase activity has just been identified in erythrocyte membranes which phosphorylates both protein 4.1 and an M_r 100,000 doublet of polypeptides in the membrane skeleton¹⁶, which could well represent the M_r 97,000/103,000 calmodulin-binding dimer¹⁴. Protein 4.1, as well as protein 4.9 and the same(?) M_r 100,000 doublet are also phosphorylated by a recently identified protein kinase C activity in the erythrocyte^{16,20}. In addition, proteins 4.1 and 4.9 are phosphorylated by a cyclic AMP-dependent kinase, but at different sites from those phosphorylated by protein kinase C²⁰.

A recent observation by Tao and co-workers²¹ provides evidence that these phosphorylations influence the organiz-

ation of the membrane skeleton. The affinity of protein 4.1 binding to spectrin *in vitro* is reduced by about fivefold on phosphorylation of protein 4.1 (but not spectrin) by a purified cyclic AMP-independent kinase from red cells²¹. It is also suggestive that the cyclic AMP phosphorylation site on protein 4.1 has just been found to reside in a spectrin-binding domain of M_r 8,000 (ref. 22). Undoubtedly, these observations are only a taste of things to come, and many of the protein associations in the erythrocyte membrane skeleton may prove to be dynamically regulated in response to physiological signals *in vivo*. □

Velia M. Fowler is in the Department of Anatomy and Cellular Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

Glaciology

Frozen news on hot events

from Claus Hammer

ICE CORES from Antarctica and Greenland provide information on the number and dates of violent volcanic eruptions and their contributions to the upper atmospheric composition of trace substances. Interpretations of the ice-core records, particularly deductions about the influence of local eruptive activity and the varying meteorological and climatological conditions during trace-substance transfer from the volcano to the ice sheets, are difficult, partly because the techniques used in various studies have not been coordinated. A recent article by R.J. Delmas *et al.* (*J. geophys. Res.* **90**, 12; 1986) concludes that the techniques used to detect volcanic signals in ice cores are all in essential agreement. Although the visible fine-ash (tephra) layers observed in antarctic ice cores must be considered separately from other data, the conclusion of Delmas *et al.*, that the sulphuric acid record in the cores reveals remote and violent eruptions, remains unchanged.

Delmas *et al.* analyse ice cores from various sites in Antarctica, both from coastal regions of high precipitation and central sites that show very low annual accumulation. They look at acidity, electrical conductivity of melted ice samples and ionic composition during the past 200 years. The lack of historic data on eruptions in the Antarctic area limits the analysis, but the authors are able to conclude that local volcanic activity in the Antarctic contributes only little to the ice-core chemistry. The fallout of fine ash is only significant close to local volcanoes (some 200 km) and the SO_2 and/or HCl from such eruptions is apparently washed out before reaching the more central parts of Antarctica, where the volcanic acid signal is par-

ticularly clear. (This conclusion does not, of course, cover violent or substantial regional activity.) Thus the major increases in sulphuric acid concentration of the central Antarctic ice is caused by remote volcanic eruptions. Transportation of the volcanic products occurred mainly via the stratosphere, although some arrived via the upper troposphere: NH_3 neutralization is virtually absent. Transfer from the site of eruption to the ice sheet during glacial times has yet to be quantitatively assessed.

The conclusions of Delmas *et al.* provide a global perspective and agree with the results obtained from the Greenland Ice Sheet (see, for example, Hammer, C.U., Clausen, H.B. & Dansgaard, W. *Nature* **288**, 230; 1980). The Antarctic ice cores provide information on past volcanic eruptions south of 20° N, whereas Greenland registers eruptions north of 20° S, so data series from the two ice sheets offer global coverage and overlapping information.

To establish a detailed and well-dated ice-eruption record; to estimate from the chemistry of the cores the influence on atmospheric composition; and to identify the geographical sites or zones of the eruptions are formidable tasks which will require not only a good deal of optimism but also new and faster techniques for analysing the details of the chemical composition of the several thousand-metre-long ice cores. More data from the smaller ice caps, for example in the Canadian arctic, are needed. □

Claus Hammer is at the Geophysical Institute, University of Copenhagen, Haraldsgade 6, 2200 Copenhagen N, Denmark.

1. Lux, S.E. *Nature* **281**, 426 (1979).
2. Cohen, C.M. *Semin. Hemat.* **20**, 141 (1983).
3. Bennett, V. A. *Rev. Biochem.* **54**, 273 (1985).
4. Marchesi, V.T. A. *Rev. Cell Biol.* **1**, 531 (1985).
5. Beaven, G.H. *et al.* *Eur. J. Cell Biol.* **36**, 299 (1985).
6. Byers, T.J. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6153 (1985).
7. Shen, B.W., Josephs, R. & Steck, T.L. *J. Cell Biol.* **102**, 997 (1986).
8. Tilney, L.G. & Detmers, P.J. *Cell Biol.* **66**, 508 (1975).
9. Stokke, B.T., Mikkelsen, A. & Elgsaeter, A. *Biophys. J.* **39**, 319 (1986).
10. Waugh, R. & Evans, E.A. *Biophys. J.* **26**, 115 (1979).
11. Shattil, S.J. & Cooper, R.A. *J. Lab. clin. Med.* **79**, 215 (1972).
12. Matsuzaki, F., Sutoh, K. & Ikai, A. *Eur. J. Cell Biol.* **39**, 153 (1986).
13. Fowler, V.M. & Bennett, V. *J. biol. Chem.* **259**, 5978 (1984).
14. Gardner, K. & Bennett, V. *J. biol. Chem.* **261**, 1339 (1986).
15. Siegel, D.L. & Branton, D. *J. Cell Biol.* **100**, 775 (1985).
16. Cohen, C.M. & Foley, S.F. *J. biol. Chem.* (in the press).
17. Ling, E. & Sapirstein, V. *Biochem. biophys. Res. Commun.* **120**, 291 (1984).
18. Raval, P.J. & Allan, E. *Biochem. J.* **232**, 43 (1985).
19. Palfrey, H.C. & Waseem, A. *J. biol. Chem.* **260**, 16021 (1985).
20. Horne, W.C., Leto, T.L. & Marchesi, V.T. *J. biol. Chem.* **260**, 9073 (1985).
21. Eder, P.S., Soong, C.-J. & Tao, M. *Biochemistry* **25**, 1764 (1986).
22. Correa, E., Leto, T.L., Speicher, D.W. & Marchesi, V.T. *J. biol. Chem.* **261**, 3310 (1986).

Chernobyl fallout on Ioannina, Greece

SIR—We have measured the γ -ray spectrum of the fallout from the Chernobyl nuclear reactor accident at Ioannina, located in northwestern Greece, nearly 1,500 km south-west of Chernobyl. We have examined air filters collected every 24 h since 29 April to monitor the γ -radiation activity, using an intrinsic germanium high-resolution detector, shielded to reduce background. The fallout probably reached Ioannina between 1 and 2 May, and the activity peaked on 5 May. After that we noticed a reduction of the air contamination, mainly due to local rains.

Figure 1 shows a typical γ -ray spectrum from one of our filters, the activity of which was measured three times at monthly intervals to estimate the half-lives of the radionuclides present. From a number of spectra, 14 different nuclides were identified by their energy spectra and half-lives and 9 additional peaks in the spectral region between 50 and 800 KeV remain unidentified.

In Table 1 the activity of each of the most prominent radioisotopes is characterized as very strong (VS), strong (S), weak (W) or very weak (VW), and the respective half-life as very long (VL), long (L) or short (S). The ^{131}I activity is defined 'strong' and its half-life as 'long'. Table 1 also lists the masses of fission decay products generated in the reactor¹. Table 2 lists the energies of the unidentified lines.

From our air-filter measurements we conclude that: (1) the relative activities for the most prominent isotopes remained unchanged during most of the period

Table 1 Radioisotopes identified from air filters

Isotope	Relative γ -activity	Half-life	Mass of fission products (mg MW ⁻¹ day ⁻¹)
⁹⁵ Nb	VW	VL	0.33
⁹⁹ Mo	W	S	107.0
¹⁰³ Ru	VS	VL	65.4
¹⁰⁶ Ru	W	VL	15.7
^{129m} Te	VW	VL	5.86
¹³² Te	S	S	
¹³¹ I	S	L	
¹³² I	S	S	
¹³⁴ Cs	S	VL	
¹³⁶ Cs	VW	L	90.4
¹³⁷ Cs	S	VL	
¹⁴⁰ Ba	W	L	38.6
¹⁴⁰ La	W	S	39.8
¹⁴¹ Ce	VW	VL	86.0

between 1 and 10 May; (2) these activities are not what would be expected from the mass of fission products and their volatility. (3) We agree with Pringle *et al.*² that the identified radioisotopes are ²³⁵U fission decay products; however, we did not locate the isotopes ⁹⁵Zr, ¹⁴⁴Ce, ¹²⁷Sb, ¹⁰⁵Rh and ¹⁴³Ce. (4) The reactor operation time based on the ratio of ¹³⁷Cs to ¹³⁴Cs activity is 15% higher than that reported by Devell *et al.*³ using the same method. (5) Finally, a trace analysis using X-ray energy spectrometry was inconclusive in locating Sr in our air filters.

The extreme conditions in the reactor core after the accident make it hard to know in what chemical compounds the radioisotopes appear in the radioactive cloud. Unfortunately, information about core conditions before and during the accident is not yet available to us, and therefore we cannot verify our assumption

Table 2 Unidentified spectral lines

Energy (keV)	Relative γ -activity	Half-life
223.4	VW	S
239.2	VW	S
257.4	VW	S
358.4	W	L
489.7	W	L
511.6	W	VL
597.9	W	L
653.9	W	S
787.6	VW	S

that the surprisingly high proportion of non-volatile nuclides (such as ¹⁰³Ru) and the absence of volatile nuclides reflect the chemical compound which hosted the nuclides at the time of release to the atmosphere.

N.G. ALEXANDROPOULOS
T. ALEXANDROPOULOU
D. ANAGNOSTOPOULOS
E. EVANGELOU
K.T. KOTSIS
I. THEODORIDOU

Department of Physics,
University of Ioannina,
453 32 Ioannina, Greece

1. Farmer, F.R. (ed), *Nuclear Reactor Safety* (Academic, New York, 1977).
2. Pringle, D.M., Vermeer, W.J. & Allen, K.W. *Nature* **321**, 569 (1986).
3. Devell, L. *et al.* *Nature* **321**, 192–193 (1986).

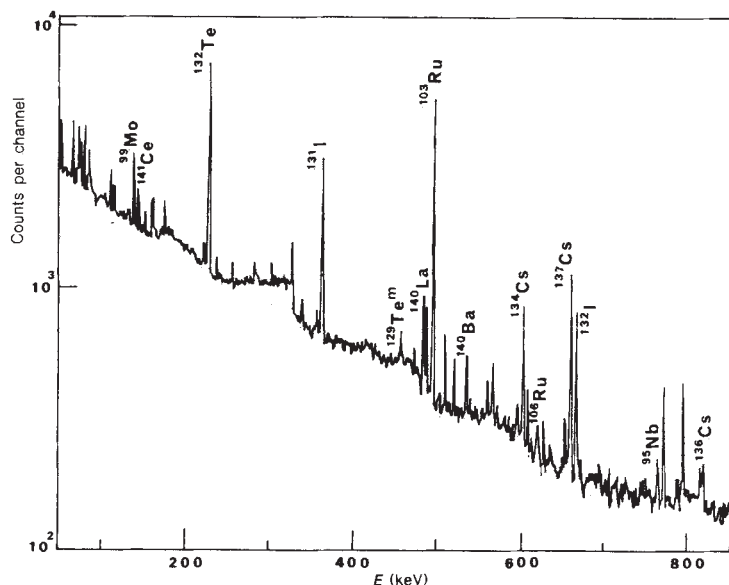


Fig. 1. A typical complex γ -ray spectrum obtained with an intrinsic Ge detector from an air filter.

The origins of the Earth's oceans

SIR—Several mechanisms have been proposed for the origin of the Earth's oceans, such as the degassing of crustal rocks over geological time¹ and the formation of impact produced atmosphere, including water, by accretion of planetesimals^{2,3} early during the Earth's formation. The available evidence suggests that sizeable oceans were present early in the Earth's history¹ and that the early Solar System

was fractionated, with nickel-iron and carbonaceous chondrites largely contained in the volume on the sunward side of Mars's orbit, and gases (hydrogen, helium, methane, water) and some carbonaceous chondrites in the volume from Jupiter's orbit outward¹.

Present theories on the origin of the oceans fail to account for the fact that Venus, a planet very similar to Earth in size and composition, has no surface water and only small amounts of atmospheric water. This could be explained by assuming that, early in its history, Earth collided with an ice moon approximately 900 to 950 miles in diameter, which could have furnished all the water presently on Earth. This collision, obviously a rare event, would have provided for rapid cooling of the Earth's crust, prevention of a runaway greenhouse effect by the absorption of most atmospheric CO₂, and rapid evolution of life. This theory is also in accord with the fractionated nature of the early Solar System.

JAY S. ROTH

Department of Molecular
and Cell Biology,
University of Connecticut,
Storrs, Connecticut 06268, USA

1. Weinaut, J.G. *Exploration of the Oceans*, (Macmillan, New York, 1979).
2. Matsui, T. & Abe, Y. *Nature* **319**, 303-305 (1986).
3. Cogley, J.G. & Henderson-Sellers, A. *Rev. Geophys. space Phys.* **22**, 131-175 (1984).
4. Walker, J.C.G. in *The Early History of the Earth* (ed. Windley, B.F.) 357 (Wiley, New York, 1976).

Frequency of dizygotic twinning

SIR—From the assumption that ovulation with intercourse leads to a live birth with probability $\frac{1}{4}$ for each egg, double ovulation may be expected to lead to dizygotic twin births with probability $\frac{1}{16}$. From that relationship, J. Diamond (*Nature* **320**, 488; 1986) suggested that an observed frequency of 4.9% dizygotic twin births among the Yoruba implies a double ovulation frequency of 16 times 4.9, or 78%.

M. Sipser (*Nature* **321**, 570; 1986) pointed out that this suggestion is incorrect. On the simplest set of assumptions, double ovulation will lead to no birth in $\frac{1}{16}$ of cases, to a single birth in $\frac{1}{16}$ of cases, and to dizygotic twins in $\frac{1}{16}$ of cases. From those relative frequencies, Sipser asserted that an observed frequency of 4.9% dizygotic twin births implies a double ovulation frequency of 7 times 4.9 or 34%.

In fact the relationship between the frequencies of double ovulation and of dizygotic twin births cannot be expressed by a constant factor. When the frequency of double ovulation is extremely low, the single births that result from double ovulation can be ignored, and the observed frequency of dizygotic twin births must be multiplied by 4 to obtain the frequency of double ovulation (single births = $\frac{1}{4}$ of

single ovulations; dizygotic twin births = $\frac{1}{16}$ of double ovulations). When the frequency of double ovulations is high, the single births that result from double ovulations must be taken into account, and the factor rises. In the extreme case, if the probability of double ovulation were 1.0, $\frac{1}{2}$ of the births would be dizygotic twins; the factor would be 7. Sipser's comment is valid only for this case, which is of no biological significance for any known human population.

The general expression to be used is $F = T/(pT + p - T)$, where T is the fraction of dizygotic twin births, p is the probability per egg of a live birth, and F is the fraction of double ovulation. When p is assumed to be 0.25, this becomes $F = 4T/(1 - 3T)$.

In any real case, the factor $[4/(1 - 3T)]$ is much closer to 4 than to 7. For the dizygotic twinning frequencies that occur in the human populations tabulated in Diamond's article (0.22% to 4.9%) the appropriate factors are between 4.03 and 4.7. In the Yoruba population the frequency of dizygotic twin births was 4.9%, and the estimated frequency of double ovulation, if $P = \frac{1}{4}$, is 23%.

DANIEL E. ATKINSON

Department of Chemistry and
Biochemistry,
University of California,
Los Angeles, California 90024, USA

How to abbreviate recombinant genes

SIR—We are all aware that recombinant DNA technologies have generated tremendous progress and spawned a literature that seems to grow exponentially. Unfortunately a custom is developing of using the abbreviation rDNA to refer to the hybrid molecules formed by uniting two or more heterologous DNA molecules; this leads to confusion, since rDNA has long been used to refer to ribosomal DNA. The abbreviation rRNA and r proteins are also in common use with r again meaning ribosomal.

To avoid confusion, I suggest the use of rt for recombinant and that editors of journals start insisting on a differentiation between rDNA and rtDNA and rRNA and rtRNA.

ROBERT C. KING

Northwestern University,
Evanston,
Illinois 60201, USA

A system of nomenclature for murine homoeo boxes

SIR—There is at present a considerable degree of confusion amongst workers on mammalian homoeo boxes because each group has tended to select a different system of nomenclature. It is now very difficult to follow the literature appropriately.

To help avoid this confusion we would like to offer the following system.

A unifying nomenclature for murine homeo boxes should consider the chromosomal location and the number of boxes on a chromosome. Clusters of boxes should preferably be numbered consecutively and in the direction of transcription. The two known clusters are probably transcribed in one orientation only and do not contain additional boxes. Thus, at least these two would be numbered consecutively from the 5' to the 3' end of the cluster. The nomenclature should further allow a logical naming of newly discovered boxes.

We suggest the new prefix 'Mox', followed by two numbers: the first one giving the chromosomal location, the second identifying an individual box on the respective chromosome. This procedure leads to the following nomenclature for a selection of published murine homeo boxes:

Proposed name	Original designation	Chromosome	Ref.
Mox 1.1	Mo-en.1	1	1
Mox 6.1	m6	6	2
Mox 6.2	m5	6	3
Mox 6.3	m2	6	3
Mox 6.4	HBT-1, MH3.	—	—
	Hox 1.3.	6	4,5
Mox 6.5	Mo 10, Hox 1.4.	6	6
Mox 6.6	—	6	—
Mox 11.1	Hox 2.4	11	7
Mox 11.2	Hox 2.3	11	7
Mox 11.3	Hox 2.2	11	7
Mox 11.4	Hox 2.1, H24.1.	—	—
	Mul	11	7-9
Mox 15.1	Hox 3, m31	15	10,11

For new isolates, we suggest the following, assuming that the description of these boxes will include the information on chromosomal location and occurrence of clusters. In this case, we propose the term 'Mox', followed by the chromosome number, a period and the lowest unused number on the respective chromosome for the most 5' box of a new cluster.

After a box has been designated in this way, the name should be kept unchanged even if new information reveals a violation of the consecutivity rule. Non-mapped isolates (from cDNAs, for example) have to obtain a preliminary designation.

PETER GRUSS

MICHAEL KESSEL

Zentrum für Molekulare Biologie,
Universität Heidelberg,
Im Neuenheimer Feld 282,
Postfach 106249,
D-6900 Heidelberg 1, FRG

1. Joyner, A.L. *et al.* *Cell* **43**, 29-37 (1985).
2. Colberg-Poley, A.M. *et al.* *Nature* **314**, 713-718 (1985).
3. Colberg-Poley, A.M. *et al.* *Cell* **43**, 39-45 (1985).
4. Wolgemuth, D.J. *et al.* *EMBO J.* **5**, 1229-1235 (1986).
5. Duboule, D. *et al.* *EMBO J.* (in the press).
6. McGinnis, W. *et al.* *Cell* **38**, 675-680 (1984).
7. Hart, C.P. *et al.* *Cell* **43**, 9-18 (1985).
8. Hauser, C.A. *et al.* *Cell* **43**, 19-28 (1985).
9. Jackson, I.J. *et al.* *Nature* **317**, 745-748 (1984).
10. Awgulewitsch, A. *et al.* *Nature* **320**, 328-335 (1985).
11. Breier, G. *et al.* *EMBO J.* (in the press).

A litany of folly

Walter Gratzer

Diamond Dealers and Feather Merchants: Tales from the Sciences.

By Irving M. Klotz. Birkhäuser:1986. Pp.120. \$24.95.

"As to the opinion which explains the putrefaction of animal substances by the presence of microscopic animalcula, it may be compared to that of a child who would explain the rapidity of the Rhine's current by attributing it to the violent movement of the numerous mill wheels at Mainz." The view is that of Justus von Liebig (echoed some time later by, I think, James Thurber, who held that wind is caused by trees waving their branches), and the child in question is Pasteur. Or try this: "Recently I pointed out that deficiencies in fundamental chemical knowledge and the lack of a good liberal education characteristic of many professors is responsible for the current decline in chemical research . . . If anyone thinks my concerns are exaggerated, he should read, if he can stomach it, a recent monograph of a Mr van't Hoff, entitled *The Arrangement of Atoms in Space*, a book swollen with infantile foolishness." That egg was laid by the great Hermann Kolbe, the first man to synthesize a organic compound, acetic acid, from its elements. In *Diamond Dealers and Feather Merchants*, Professor Klotz has assembled a wonderful collection of such philippics in his exploration of how these thunderous gaffes come to be perpetrated — hostages heedlessly offered to fortune by those best placed to know better.

Perhaps success, like power, tends to corrupt; Ambrose Bierce in *The Devil's Dictionary*, ever a sound guide to human nature, encapsulates the matter thus: "Intolerance is natural and logical, for in

"Klotz suggests that the strange passions that cloud the judgement and take the reason prisoner can express themselves as much in credulity as in obduracy."

every dissenting opinion lies an assumption of superior wisdom." The uninhibited ferocity and vindictiveness with which academic vendettas were publicly conducted belong of course to a bygone age of polemic. The heavy, snarling sarcasm, the pseudonymous lampoon on the rival's work, have given way to a little discreet disparagement in the intimacy of the grant committee; but the results are similar and it is probably an inalienable part of the scientific process. Max Planck said that radical scientific innovations do not prevail by overcoming prejudice, but rather through the eventual demise of their

opponents, to make way for a new generation of scientists whose prejudices have not yet hardened.

Klotz suggests that the strange passions that cloud the judgement and take the reason prisoner can express themselves as much in credulity as in obduracy. Chauvinism and political engagement play a part in two spectacular scientific debacles that he chronicles with relish, but not without compassion: N-rays were named after the city of Nancy by the respected French physicist, Blondlot. Klotz surmises that Blondlot's formative years would have been darkened by the defeat and disintegration of the Second Empire and the loss of much of Lorraine, which brought the German border almost to Nancy. This must have inflamed the rivalry, already intense, between French and German science. In addition, Blondlot was probably still grieving at the turn of the century that the discovery of X-rays had eluded him.

N-rays were an electromagnetic radiation emitted from electric discharges and many other sources, such as hot metals. They were taken up with fervour by French physicists, chemists and indeed biologists, for it soon emerged that nerves, muscles and even enzymes were abundant emitters. By 1904 the ratio of papers in the *Comptes Rendus* on N-rays to those on X-rays was 53:3. The outside world was more cautious and the end came when the great American spectroscopist, *farceur* and eccentric, R.W. Wood (the man who trained his cat to clear the cobwebs from the 40-foot optical path of his spectrograph), visited Blondlot's laboratory to observe the measurement of wavelengths in the N-ray spectrum, dispersed by an aluminium prism. The room being dark, Wood adroitly trousered the prism, with no detriment to the process of data collection. Wood spared no feelings when he subsequently described the events in *Nature*. J.A. Le Bel expressed what many French scientists must have felt: "What a spectacle for French science that one of its distinguished *savants* determines the position of lines in the spectrum while the prism sits in the pocket of his American colleague."

Klotz seeks a parallel between the N-ray affair and the curious mirage, half a century later, of polywater — an associated form of water reported by the Russian surface physicist, Boris Derjaguin, to form in fine silica capillaries — for

here too political prejudice may have exerted a baleful influence. J.D. Bernal, who was preoccupied at this time with the structure of water, must, so Klotz believes, have badly wanted Soviet science to come up trumps. (He continued, for instance, to support Lysenko long after the truth had come out in all its squalor.) Bernal did in the event jump off the polywater bandwagon before that vehicle careered over the precipice, taking many reputations with it. Joel Hildebrand, the Methuselah of physical chemistry who published his last paper at the age of a hundred, seemed in no doubt from the outset. "Water and silica," he wrote, "have been in intimate contact in vast amounts for millions of years; it is hard to understand why any ordinary water should be left." Another physical chemist, R.E. Davis, put it succinctly: "We must conclude that all polywater is polycrap and that the American scientists have been wasting their time studying the subject". In retrospect it seems that most of these scientists could have believed in polywater only because they wanted

"Chauvinism and political engagement play a part in two spectacular scientific debacles that Klotz chronicles with relish, but not without compassion."

polywater to exist. Derjaguin accepted earlier than many that he had been wrong, and emerged with dignity from the reeking shambles.

Klotz concludes with some examples of credulity and ignorance to make your toes curl. Thus: "From experience with non-science college students, I know that the majority predict that if a silver dollar and a silver dime are simultaneously dropped from the top of the Sears Tower the former will hit the ground first. An even larger fraction when asked whether a two-inch line or a one-inch line has more points will vote in favour of the former: the remainder exercise their democratic privilege of not voting. The general public would overwhelmingly endorse [this] view." I understand the flat-Earthers are still among us, satellite pictures notwithstanding: common sense lives.

Professor Klotz's name is familiar to all protein chemists, and generations of science students were brought up on his celebrated textbook of chemical thermodynamics. He now reveals himself as a writer of wit, wisdom and originality in a different genre. This admirable little book will give much pleasure. □

Walter Gratzer is in the Medical Research Council's Cell Biophysics Unit, King's College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

Looking beyond creature comforts

Marian Stamp Dawkins

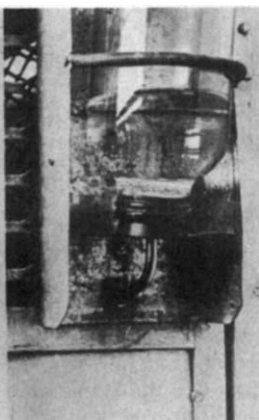
Laboratory Animal Husbandry: Ethology, Welfare and Experimental Variables. By Michael W. Fox. *State University of New York Press*: 1986. Pp.267. Hbk \$39.50; pbk \$9.95.

I MUST confess that most books on animal welfare give me a sinking feeling even before I read them. I have a prejudice, no doubt grossly exaggerated and unrelated to the real world, that there are only two types of book about this subject. The first treats animals in a mystic, reverential way and castigates humans for interfering in the natural balance of things and, sometimes, for having anything to do with animals at all. The other goes to the opposite extreme and talks mostly about sq cm of cage size and the regular collection of faeces. There seems to be little in between: it is as if one has either to be poetical in one's attitude to animals or adopt the earthiest sort of banality about the size of their cages or the cleanliness of their drinking water.

The sinking feeling occurred when I first looked at *Laboratory Animal Husbandry* by Michael W. Fox. However, despite its title, the book certainly does not fall into either of the usual categories. On the contrary, it is innovative in several important ways.

First, it documents recent evidence about the effects that totally unexpected factors may have on animal health and well-being. It is perhaps not surprising that noise, ventilation and ambient temperature can have significant effects on the health of animals and their susceptibility to disease. But it is probably less well known that the exact age of weaning and the amount of handling from a human being can affect, among other things, plasma corticosterone, prolactin and growth hormone levels. The behavioural and physiological changes that have been found can in turn affect not only the animals' own well-being but the results of experiments on, say, drugs. Some of these unexpected factors may radically alter the conclusions about how effective a treatment is.

A second, related point that Fox makes concerns an animal's social environment. It should be obvious, but it clearly is not to judge from current practice, that an animal's welfare cannot be safeguarded simply by concentrating on the bare necessities of food and water or even on the physical environment of its cage. Many animals are highly social. They interact, both positively and negatively, with other members of their species; it should not be



Improved conditions for animals could benefit humans too

Camera Press

surprising, therefore, that disturbances in their social lives can significantly affect their behaviour and health. Fox argues that our understanding of human ailments is far from complete and it is often impossible to pinpoint why one person becomes ill when another does not. If we are attempting to use animals as models of puzzling human diseases, he argues, we will continue to be ignorant unless we also begin to consider the complexity of the factors, including social ones, that affect the development of diseases in animals.

He concludes that there is a double imperative on us to improve the welfare of laboratory animals. One is for the animals themselves: if we use them we should look after their welfare better than we do now by taking into account the range of factors, both physical and social, that can affect them. The other is for humans: if we want good scientific results, we should be more careful about the conditions in which we keep animals. There follows a useful chapter giving practical suggestions as to how laboratory animal husbandry could be improved for the benefit of both human and non-human animals. Another chapter contains a review of the main findings on pain and suffering in research animals.

One particularly appealing feature of the book is that the discussion of animal rights is kept to the last chapter. This means that even if you disagree with Fox on this issue or lose sympathy with the "land ethic" or "One Medicine", the rest of the book with its more down-to-earth reviews and suggestions, loses none of its value. The ethical issues are discussed but they do not intrude, so that the book can potentially appeal to, and be profitably read by, people with very divergent attitudes to animals.

The common thread is a concern for the welfare of animals in laboratories. Fox's contribution is to bring common-sense and research findings to bear on the problem of how we might improve it. He points to the arbitrariness of the present guidelines and laws as to how animals should be kept (the fact that recommended cage sizes for dogs, cats and primates in Britain are double those suggested by the

American authorities, for instance). And his insistence that we should take more notice of the social needs and interactions of animals, and not just think in terms of their physical comfort, is a message that cannot be broadcast too loudly or too frequently. □

Marian Stamp Dawkins is Fellow in Biological Sciences at Somerville College, Oxford OX2 6HD, UK. She is also author of *Animal Suffering: the Science of Animal Welfare* (Chapman & Hall, 1980).

Economical images

Peter Hawkes

Transform Coding of Images. By R.J. Clarke. *Academic*: 1985. Pp.432. Hbk. \$69.50, £60.50; Pbk. \$34.95, £31.

FROM the point of view of a computer, an image is a voluminous and cumbersome object. The human eye and brain can handle a continuous-tone picture effortlessly and almost instantaneously; for a computer, the image must be broken into small zones, each of which corresponds to a number representing the grey-level. There are typically 256 of the latter, and the number of image-elements or "pixels" can easily reach a million. For these values the computer already has 8 megabits to cope with, but if the image is coloured rather than black-and-white this must be multiplied by three. Worse still, although we may regard such an image as a 1024×1024 matrix, each element of which is an (eight-bit) number representing the grey-level of the corresponding pixel, in image processing we are also interested in statistical properties. If the image is represented by an $N \times N$ matrix ($N = 1024$ in our example), then the matrix corresponding to such statistics will be $N^2 \times N^2$.

Everyday images contain much redundant information — uniform or slowly varying areas for example — and it is natural to ask whether a typical image could not be represented in the computer in a more economical way. This question has long preoccupied the image proces-

sing community and, although not quite so pressing with the advent of parallel and special-purpose computers, it will surely always be important for archival purposes. Various ways of coding images that occupy less memory space have been explored, and Roger Clarke's book is a detailed introductory account of one of these ways, notably the use of transforms, of which the Fourier is the best known.

The underlying idea is simple. Fourier transforms tell us what weight is associated in some signal with the corresponding frequency. We can therefore expect that if we suppress frequencies with very low weights, and reconstruct the signal from the remainder, the result will differ only slightly from the original. Converting this qualitative notion into quantitative terms is not so simple: how low must a weight be to be regarded as very low and hence negligible? How slight must the difference between original and reconstruction be to be acceptable? Is the Fourier transform a good choice and if not, which is the best transform to use? It is these and related questions that Clarke sets out to answer. He does so in eight chapters, plus an introduction, in which he examines image statistics, suitable orthogonal transforms, quantization, practical coding, the human visual response, fast transforms, errors and noise and, finally, rate distortion theory and coding. Seven appendices are concerned with specific mathematical points.

How successful he is depends on the sophistication of the reader. Everything is spelt out in detail: the laboriously gentle introduction of a matrix product at the beginning of the chapter on orthogonal transforms is a glaring example. Any physicist who has followed a course on quantum mechanics, not to say applied mathematics, will be exasperated by this, as will many electrical engineers, especially those with any familiarity with coding. Conversely, many readers will appreciate the detailed and careful definition of the many matrices that arise in image coding, and will be grateful for the frequent numerical examples that add flesh to the bare mathematical bones.

It must be said too that Clarke's is a leisurely style: "At the outset, it is worth clarifying one matter concerning the application of the theory . . .", he observes; and he likes telling us that "it is a truism nowadays to say" and that "it is no exaggeration to say that . . .". His discursiveness seems at odds with the aim of transform coding! Nevertheless, he has written a clear, readable and easy introduction to his subject, and those of us with images to code will often turn to his book to see what he has to say about some specific problem. □

Peter Hawkes is Directeur de Recherche at the Laboratoire d'Optique Electronique du CNRS, BP 4347, F-31055, Toulouse Cedex, France.

Catalogue of inequities

John Maddox

The World of Science and the Rule of Law. By John Ziman, Paul Sieghart and John Humphrey. *Oxford University Press: 1986. Pp. 343. £19.50, \$37.*

This is not so much a book about the world of science and the rule of law as about what are called, in the United States, human rights — conflicts between personal liberty and the interests of the state. John Ziman, Paul Sieghart and J. H. Humphrey are, respectively, distinguished physicists, lawyers and biologists whose interests have repeatedly strayed beyond the narrow boundaries of their professions. Their failure, in this book, to do much more to describe a haunting problem is not their fault, but a measure of its difficulty.

The book has its origins in a pamphlet published more than ten years ago by the British organization called the Council for Science and Society, an unpaid and sometimes unprofessional Office of Technology Assessment. The starting point is the observation that the complaints of scientists that their professional freedom is too often administratively constrained is often at odds with the way in which most formal constitutions formally guarantee the freedom of the individual to do what he pleases.

Given the signature of the agreements on European security at Helsinki in 1979, the book cannot but in part be an account of the lamentable record of what has happened since. But little has changed. People who work in states where freedom of communication is guaranteed by the laws still discover that their letters are opened, sometimes to be confiscated, and their telephones disconnected. People whose governments are supposedly bound by the rule that employment should be open to all still find themselves denied the chance to work at what they do best when they give offence; the case of the Jewish refuseniks in the Soviet Union is familiar, and necessarily a large part of the litany of oppression making up this book. Indeed, the Eastern bloc gets a real drubbing.

The authors might have made more of the way in which many western European States have been compelled by the European Convention on Human Rights to do the decent thing, against their first inclinations. Maybe there is more freedom now than, say, in the eighteenth century, even though the continuing deficiency of what is on offer remains just as offensive now as then. So the authors are right to ask why performance should so often fall short of promise. Why? A discussion of this point

would have been rewarding.

The truth is that the human failing of believing well and behaving badly applies to governments as well. Let us hope they may improve. How might the professional community urge them on? After what is a moving catalogue of injustice, the authors have next to nothing to offer by way of remedy. They approve of the way that two successive presidents of the Royal Society have "spoken out", but have surprisingly little to say about the need for the outspoken to contain their protests within the other fellow's law. UNESCO's statements on the status of scientific workers are approvingly reprinted as an appendix, but the authors seem to share the general conviction that the source of all the words has no act to perform: ICSU is a better bet. Weiskopf's dilemma, that an illiberal society ostracized into exclusion can only become even more illiberal, is described but not resolved. And nowhere in this flat peroration is there a clear explanation, for the benefit perhaps of the young and cynical, of why it is that, if the law has a built-in momentum sweeping illiberality out of the way, then surely science must work the same magic but more powerfully. But is that not the only way in which it will come out? □

John Maddox is editor of Nature

NEW TITLES BY ANS

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

"Introductory Nuclear Reactor Statics"

by *Karl O. Ott*
Winfred A. Bezella

Recommended as an undergraduate or graduate textbook. Valuable source book for reactor analysis and design.

"Introductory Nuclear Reactor Dynamics"

by *Karl O. Ott*
Robert J. Neuhold

Examines fundamental principles of Nuclear Dynamics. Companion volume to Nuclear Reactor Statics, published in 1983.



For additional information write.

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Reader Service No.7

Plenum: The Practical Source

MICRODOMAINS IN POLYMER SOLUTIONS

edited by Paul Dubin

Brings together current research and experimental studies on concentrated solutions in both synthetic and natural polymer systems. Volume 30 in Polymer Science and Technology.

0-306-42110-0/proceedings
472 pp./ill./1985 \$75.00
(\$90.00 outside US & Canada)

Forthcoming

MULTINUCLEAR NMR

edited by Joan Mason

Covers the entire NMR Periodic Table from a chemical and biochemical standpoint. It can be used as a handbook of first resort by the professional or as an advanced textbook for students in chemistry or biochemistry.

0-306-42153-4/approx. 750 pp./ill.
1987

THE JOURNAL OF SCIENTIFIC COMPUTING

Editor-in-Chief:

Steven A. Orszag

This new journal serves as a forum for state-of-the-art developments in scientific computing over broad interdisciplinary grounds. The key goals of the journal are to stimulate the development and dissemination of ideas related to large-scale scientific computing and to foster increased communication between university researchers and the designers of computer hardware and software.

Volume 1, 1986 (2 issues)
Institutional rate:
\$50.00 in US/\$56.00 elsewhere
Personal rate:
\$25.00 in US/\$29.00 elsewhere
Volume 2, 1987 (4 issues)
Institutional rate:
\$95.00 in US/\$107.00 elsewhere
Personal rate:
\$35.00 in US/\$41.00 elsewhere

CHEMISTRY BY COMPUTER

An Overview of the Applications of Computers in Chemistry
by Stephen Wilson

In this broad survey, Dr. Wilson covers the methods used by the computational chemist, including quantum chemistry, molecular mechanics, collision theory, chemical kinetics, organic synthesis, Monte Carlo simulation, and molecular dynamics.

0-306-42152-6/224 pp. + index/ill.
1986/\$37.50 (\$45.00 outside US & Canada)

Plenum Publishing Corporation
233 Spring Street
New York, N.Y. 10013-1578
In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Fifty years on

Warwick Bray

American Archaeology Past and Future: A Celebration of the Society for American Archaeology 1935-1985. Edited by David J. Meltzer, Don D. Fowler and Jeremy A. Sabloff. *Smithsonian Institution Press: 1986. Pp. 479. Hbk \$35, £35; pbk \$19.95, £19.75.*

ON 28 December 1934 the Society for American Archaeology came into existence. The following year the initial volume of *American Antiquity* appeared and the first annual meeting of the SAA was held (with 8 papers and 75 participants). *American Antiquity* has since become the leading journal in its field, and many people would argue that the SAA meeting is the major conference of the archaeological year. The present volume is a celebration of the society's 50th birthday — an appropriate age for looking back to the past, taking stock of the present and trying to set the younger generation on the right path for the future.

The 17 contributions are by North American practitioners belonging to the older and middle generations, but representing a range of differing archaeological theologies. They are neither complacent nor self-congratulatory. Cumulatively, the book adds up to a lively, and conspicuously well written, critique of the strengths and weaknesses of archaeology in North America. By sheer weight of numbers, the Americans have played a dominant role in changing the theoretical basis of archaeology (in the non-Communist world, at least), so this study is of more than parochial importance.

The editors divide the volume into three sections. The first looks backwards: it includes a personal reminiscence by Jesse Jennings, and historical surveys of the development of field techniques, the rise of the conservation ethic and the changing employment of the concept of culture. Bruce Trigger examines archaeology as a social phenomenon which reflects the changing interests and prejudices of the society that sponsors and pays for it (as an example, he notes how today's explanatory models reflect "current anxieties in middle-class American society about unchecked population growth, environmental destruction and the depletion of nonrenewable resources"). In similar vein, Don Fowler observes that the roots of the majority of Americans do not lie in the Indian past and, in consequence, there is no strong impulse to identify with and conserve that past. Each generation, it seems, gets the archaeology it deserves.

The middle section of the book consists of state-of-the-art surveys of three major problems: the nature of hunter-gatherer

societies, the origins of food production, and the evolution of civilized states and empires. These papers, though good of their kind, do not treat the issues historically and therefore seriously interrupt the continuity of the book. The main theme returns in the final section on current trends and future prospects (the role of mathematics and formal methods, problems of cultural resource management, how to reconstruct the symbolic and cognitive values of extinct communities, and an evaluation of our long-term intellectual options).

Sections One and Three should be compulsory reading for any one, not only Americanists, interested in the historical development and present health of the discipline. As the editors point out, in 1935 American archaeology was remarkably uniform, with generally agreed goals (historical reconstruction, above all) and universally accepted research methods. That is not true today, and one of the book's many virtues is that the polemics of recent years are not merely described, but also evaluated from a long-term historical perspective. One of the faults of recent Messiahs and their followers is a reluctance to read anything more than 10 years old, and it is good to see Donald Grayson demonstrating that "middle range research" — though not under that jargon title — goes back to the days of the controversy over eoliths, and to read Jennings's casual comment that he had "discovered, studied and discarded" (my italics) the philosophers of science long before Hempel was so uncritically taken up by the New Archaeology. Some scholars, including some of the most innovative, are traditionalists by considered conviction, not through ignorance or lack of imagination.

Without naming names or reopening old wounds, one can recognize from the data in this volume a fairly standard cycle of events. A new movement emerges (generally based on ideas developed in another discipline, and often just as these are beginning to go out of fashion on their home ground), is accepted uncritically by its devotees, fails to deliver the promised goods, gradually loses its adherents and is in turn replaced by another "new wave". The historical view presented in this volume demonstrates that, in the long run, the good (and there is always some) is retained and adopted into the mainstream of archaeology, and the sillinesses — like puberty spots — fade away with maturity. If archaeology is not to stagnate, this kind of intellectual ferment is vital to the well-being of the subject. Like most of the contributors, I look forward to the future with a guarded optimism. □

Warwick Bray is Reader in Latin American Archaeology at the Institute of Archaeology, University of London, 31-34 Gordon Square, London WC1H 0PY, UK.

Breeders' rights and patenting life forms

from Jean-Pierre Berland and Richard Lewontin

Academic freedom and scientific advance could be the casualties if the patenting of microorganisms becomes the norm.

WITH the advent of genetic engineering, biologists have suddenly become acutely conscious that plants and animals can be factors in the production of consumer goods: like machines, they are commodities, sold for profit to manufacturers. This development has raised a host of issues about property rights in living organisms, the patenting of life forms and other questions that arise when commodities are produced for profit.

The excitement over genetic engineering may, however, obscure the fact that plants and animals have been factors of production of consumer goods for a very long time. The production of seeds and of animal breeding stocks for sale to farmers by seed companies and breeders have long ago raised the legal and economic issues of life forms as private property, used by manufacturing firms to appropriate profit. Moreover, research devoted to the production of new varieties or to methods of plant and animal breeding has always been, at least in part, influenced by the drive to increase the profits of breeders and seed companies. An analysis of the economic forces in the seed industry, as they have operated in the past and at present, may provide us with a framework for understanding future questions raised by newer biotechnologies.

Seed business

Seeds are a special kind of factor of production because, at least potentially, they are reproduced in the production process itself. Thus, in principle, a farmer could produce his own seed by withdrawing a small portion of his crop from the market. The problem for the seed company is to convince the farmer not to do this, but to buy this factor of production each year, just as he does fertilizer and herbicide.

There are two pathways open to the seed producer. The first is to achieve economies of scale, that is, to produce seed of the appropriate quality cheaper than can the farmer himself. As we will show, the profits from economies of scale are, in general, quite small. The second is to reap monopoly profits by creating seeds that are consumed in the production process; by providing objects that are not really seed in the biological sense that they are self-reproducing.

This is the secret of the hybrid method of breeding, which makes it impossible for

the farmer to hold back some of his crop as seed for the next year and which forces him to return each time to the seed company. Such monopoly profits can be very high. The search for them has had a strong influence on the direction of breeding research and on the drafting of laws and regulations governing seed production and sales in North America and Europe for 60 years.

It is our purpose here to provide a sketch of the economic analysis of the seed industry and to show how the attempt to maximize profits in the seed industry has influenced breeding and its organization. We will then discuss briefly the consequences of breeders' rights and the essential contradiction involved in patenting life forms.

Property rights

Any bag of seeds has two characteristics that must be distinguished carefully. The first set of characteristics, that the seeds should be weed-free, sized to fit mechanical planters, free from seed transmitted diseases, that they should germinate consistently to reach the proper plant density and that they should be available when needed in convenient forms at a reasonable price, are all traits that are a function of the technical process of seed production and marketing; they contribute to the profits of the seed producer under the heading of economies of scale. The second set pertains to the variety itself, its adaptation to the environment, its earliness, disease resistance, stress tolerance, yield, taste, processing qualities, nutritive value and so on. These characteristics are the result of the breeder's work, whether public or private. They are embodied in the genetic framework of the variety, are of a strategic nature and have an immense value. Marquis wheat, for instance, because of its earliness and disease resistance, made it possible to expand vastly the wheat-belt in the northern United States and Canada.

Yet, breeders find themselves in dire difficulties, for they have no property rights in the result of their labour. Farmers, no matter how much they benefit from the breeder's work, will pay only for the services of the seed producer, but not willingly for the work of the breeder. This is why breeding has had to be publicly financed until varieties — the result of breeding research and the invention that

stems from it — become saleable commodities.

Historically, this has been done in two different ways. In Europe, governments have undertaken regulatory activities for a very long time and, more recently, this long-standing administrative protection has been reinforced by a legal system of plant breeders' rights. In the United States since the mid-1930s, the same goals have been achieved in the production of F_1 hybrids, as is the case with corn. We will describe the historical origins of the system of breeders' rights and analyse its anticipated consequences before we turn to the biological solution sought by US plant breeders in the 1910s and 1920s.

Breeders' rights

Breeders' rights have existed in Europe for more than 50 years, thanks to seed trade regulations, although the final convention forming the *Union pour la Protection des Obtentions Végétales* (UPOV) was not signed until 1961 and not ratified by national legislation until 1970. To understand how these rights were finally developed, we have to go back to the beginning of the twentieth century.

The first attempt to create a system of protection was to extend the patent system to new varieties, but patent lawyers rejected this solution on the ground that one of the basic requirements of patent law cannot be satisfied with biological material. Specifically, even full disclosure of the breeding work would not make it possible for anyone to reproduce the variety. Accordingly in France, breeders had to find more imaginative solutions, which resulted in a series of decrees, beginning in 1922, requiring that (1) seeds be sold under their proper variety name, labelled by variety and producer; (2) the breeder of the variety becomes the owner of the variety name, which is registered; and (3) only seeds of registered varieties can be offered for sale. Therefore, the breeder owns the variety.

In 1928, an "identity and purity" service was set up within the French ministry of agriculture to check the identity of the genetic material offered for sale and to control the purity of the seeds, thus establishing a *de facto* protection of breeders' rights.

These are the two pillars of the European system of protection: — the *catalogue*,

originally called the register, and (borrowed from the Dutch) *certification*, to protect farmers from poor plants. Fields are inspected, tests are run and plants are sold with a certification tag attached to the bag. If the tags are numbered and mention the name of the variety and of the producer of the plant, and if the records are kept, the breeder knows exactly the quantities of his variety sold.

When the catalogue and certification coexist, breeders are protected and *de facto* property rights to plant varieties are created without the intervention of legislation (see ref. 1). Only in the 1970s did European countries adopt legislation on plant breeders' rights giving this regulatory procedure a legal basis. But no such system can provide a mechanism to set the price of a breeder's work, nor is there a market mechanism that can establish one. This problem had to be solved by a regulatory mechanism to decide what the breeder's royalty should be.

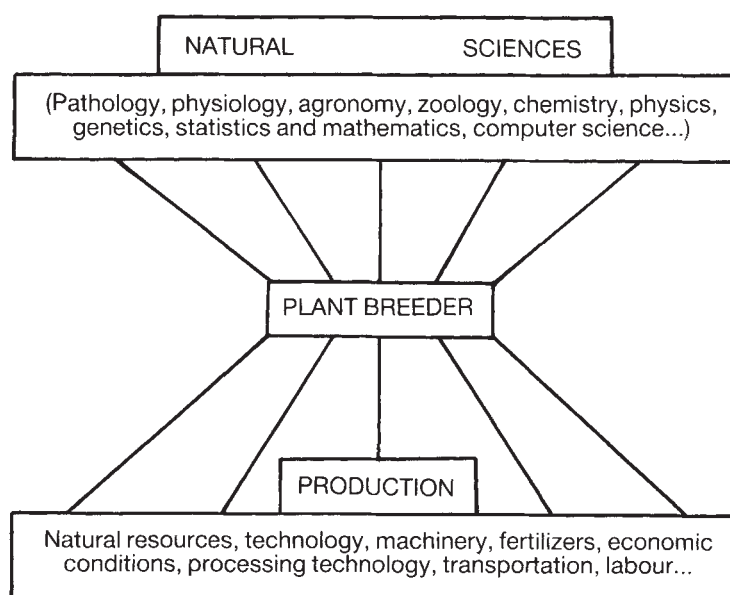
For clarity, we shall assume that the two functions of breeding on the one hand and of commercial seed production on the other are separate. We might suppose that, under the regulatory system of protection (or a breeders' rights system), the breeder can collect his royalties. The seed company can then add the royalty to its costs to set the price of the commercial seed, so that the cost of the breeding work would be paid for by the farmer.

This is not the case, however. Why would a rational farmer pay a second time for something he has already bought and still possesses in the form of his seed crop? The harvested grain has to be processed into seed; the farmer has the choice of doing it himself or buying the seed from a commercial seed company. Whether he chooses one or the other hinges on the importance of the economies of scale in seed processing.

Cleaning and sorting the harvested grain to get rid of foreign matter and of weed seeds, coating the seed with fungicides and insecticides and storing and keeping it in good condition until the next growing season, after checking that it germinates, are simple and cheap operations; in France, they represent about 30 per cent of the value of the harvested seed. The sorting and cleaning equipment is cheap, easy to operate and readily available from country elevator operators at off-season times and prices. If necessary, machinery can be purchased and used cooperatively.

All this means that the profitability of the seed companies is low. Requiring them to pay for the breeding work is, for many, the last straw. In Britain, for example, a number of small seed companies went under after the breeders' rights law was passed. In France, where commercial seed is generally produced by cooperatives on behalf of their members, this

The strategic position of plant breeding



operation is often in the red.

To further illustrate this point, in 1985 the French government made it illegal for seed companies to process farmers' grain into seed. Why had seed companies apparently been undermining their own market? Because they were thus able to keep for themselves the benefit of economies of scale when processing farmers' seed, whereas they had to pay the breeder his royalty for every bushel of commercial seed they sold. Breeders and breeding companies sought to close this loophole and succeeded.

In the United States, the law allows farmers to use the harvested crop to provide seed for the next generation and even to sell it (although advertising is forbidden). This law is ambiguous on the question whether farmers may collectively use equipment to process their grain into seed, or have it processed for a fee, or trade the seeds of various varieties within a clearing house to avoid having to grow, harvest and process small quantities of each cultivar. It is likely that strong pressure will be exerted in the future by breeders and private seed companies to prohibit such operations, as the example of France demonstrates.

From this analysis, it can be seen that it does not make sense for a business enterprise to seek to integrate breeding and seed production: one loses what the other gains. (There is a striking difference with hybrids.) But breeding is a strategic activity that benefits the entire society. Should the cost be met by taxpayers or by a small group of enterprises, the small seed companies? And if the second, what steps can be taken to ensure that seed producers' profits will be sufficient to sustain such a

crucial activity? Especially now that breeding autogamous plants is less and less of a craft, and has become a specialized and capital-intensive operation, assurance is not easily arrived at.

As well as the hybrids, there is another significant exception to the qualitative conclusion of this analysis — the case of the occasional crops, such as soya beans, where the economies of scale in seed preparations can be very great. Because of the supposed fragility of the soya bean seed, it is said to be difficult for farmers to reproduce their own seed from the harvested crop. The question why mechanization research at agricultural experimental stations has not devoted more attention to harvesting techniques that would be less damaging to soya bean germ apart, soya bean is idiosyncratically closer to the category of monopoly profit seed production (see below) than to that of the small-grain cereals.

Changing rules

When no breeders' rights are in force, a public breeder will replace a variety only if he has developed a clearly better material. If a variety is widely used for years on end, so much the better. His fame, and that of his experiment station will be increased, and the farming community will lobby for increased appropriations when it becomes necessary.

For a private breeder, the problem is quite different. He has first to make his company profitable, which requires that there should be a commodity for sale. The strongest sales pitch is to offer "new" varieties. In other words, a private breeder has a vested interest in reducing as far as possible the lifetime of his varieties, to-

wards the ideal that farmers should adopt new varieties every year.

It is not therefore surprising that the statistics presented to the US Congress in 1980 have been described² as "impressive". The same source says that "... three to six times more varieties of wheat, soybeans, and cotton were produced in the decade after the passage of the 1970 Act than in the preceding decade ... This is probably a case in which the patent system actually achieves the intended goal of increasing research investment."

No seed company will advertise a non-proprietary item. Product differentiation through the creation and sale of proprietary varieties is essential if it is to remain competitive. The increase in the number of varieties offered for sale after the Breeder's Rights Act was passed in the United States is, for these reasons, not an adequate measure of the breeding effort aimed at increasing the agronomic value of the genetic material because, meanwhile, the objective of the breeders has shifted from that of increasing farm productivity to that of giving sales arguments to the marketing departments of the seed companies.

Most breeders recognize that varieties offered by private companies are very similar². The argument that the number of varieties offered for sale has been multiplied, implicitly suggesting that it could be an index of the actual breeding work, is tantamount to measuring a temperature first with a Celsius thermometer, then with a Fahrenheit thermometer, and concluding that the temperature has approximately doubled. What the figures show is that breeders' rights increase oligopolistic non-price competition, but they leave unsettled how, and what sort of agronomic progress flows from the new rules of the game and whether it is more or less efficient than the alternative of publicly funded breeding work.

We would add at this stage that to contrast "market" and "public" breeding work is simplistic and inappropriate. The expensive, risky, time-consuming and unattributable task of developing new breeding techniques is carried out by the "public sector" with the intention that they should become the general property of all seed companies. At this point, private breeders step in to add small touches that can be made proprietary at small expense.

The structure that replaces a totally public breeding system is thus one of division of labour between public research, which does the non-proprietary work, and private breeders, who endeavour to use the work of public breeders in their own breeding programmes and proprietary commercial varieties. Far from disappearing into oblivion, public research will remain active, but will have to give up the production of commercial varieties. The

process of adjustment may be long and arduous, especially for the older public breeders who may cling to their ideal of serving the farmer, whereas they have now first to serve the seed industry, hoping that the "invisible hand" will accomplish what they used to do consciously. The United States and France, where it is a quasi-official policy that public research should keep out of commercial varieties, clearly illustrate the division of tasks created by a breeders' rights system.

Biological implications

It took four years for a committee of experts to elaborate the statute of UPOV (the Union for the Protection of Plant Invention). One of the stumbling blocks was the definition of the object of the convention¹:

The word variety ... applies to any cultivar, clone, line, strain hybrid susceptible to be grown and satisfying the dispositions of paragraphs (c) and (d) of Article 6. These paragraphs state that a new variety must be sufficiently homogenous and stable.

In fact, as UPOV itself recognizes, a variety is almost impossible to define³: "... the word 'variety' which is currently used does not have a commonly accepted precise definition". Since the very object of the convention is without a biological definition, we will not use the word "variety", which refers to a common sense notion and may confuse the issue, but instead use a meaningless term, *varitrick*, which is legally defined.

A new *varitrick* must be distinct from others. It has to be homogenous — the plants belonging to the same *varitrick* must look alike — and it has to be stable — a *varitrick* should remain identical or almost identical over time. While these are the minimum criteria of patentability, they are also nothing other than a description of the steps of the method of breeding widely used for self-pollinated plants or to produce inbreds. Breeders artificially cross homozygous plants which display desirable qualities and then breed a series of generations, sorting out the plants which display the qualities of both parents, until they get a homozygous plant which, by definition, is a new (distinct) homogenous and stable *varitrick*. We thus have the principles of *distinctiveness*, *homogeneity* and *stability* (D-H-S) which are the basic principles for protection of breeders' rights all over the world.

In essence reliance on D-H-S as a criterion for a *varitrick* to be the property of a breeder outlaws other methods of breeding. The biological object, the variety, with its immense potential of variation, adaptation and evolution, is reduced to a *varitrick*, a much narrower legal object. For the sake of establishing property rights, plant breeders' imagination has to be channelled into this narrow definition:

time, energy, and money have to be expended in fixing small anatomical differences (sometimes sensitive to the environment) in order to comply with the legal definition.

Let us again stress the self-contradiction of the system: if a *varitrick* is stable, the harvested grain is genetically identical with the seed and there is no need for an integrated seed industry. Breeders cannot make a living out of their work if they do not squeeze it out of the seed producer. If the variety is unstable and if the breeder/seed producer knows how to control this instability and to maintain the basic agronomic characteristics of his variety (by F_1 hybrids, for example), only then will the farmer be ready to pay dearly for the breeders' services and be compelled to pay for the expenses of research and development.

Hybrids and monopoly

Barton, in his review of breeders' rights raises the science-fiction possibility that "DNA engineering will be applied to make the second generation of a seed artificially sterile ..." and concludes that such an 'innate patent system' could pose enormous social costs in a concentrated industry. This forecast is to the point but at least 60 years behind the times. The F_2 generation of hybrid maize, if not biologically sterile, is economically unusable as seed, producing anywhere from 20 per cent to 40 per cent less than the F_1 hybrid. For all practical purposes, such a loss of yield amounts to biological sterility.

In good maize areas, the use of F_2 seeds instead of commercial hybrids will cause a loss of 45 bushels per acre if a potential yield is 150 bushels per acre. The maximum price per bushel that an economically rational farmer will be willing to pay for his hybrid seed corn is theoretically equal to the loss per acre (45 times the price of maize per bushel) multiplied by the area sown with a bushel of seed corn (about 4 acres), which is the equivalent of 180 bushels of maize.

Even at present low grain prices, a bushel of hybrid seed corn is thus a potential gold mine, although seed companies do not, in fact, realize the maximum price. With grain selling at about \$3.00 per bushel, the maximum price of hybrid seed would be approximately \$450 per bushel, while the price, if the farmer could produce his own non-hybrid seed, would be a little over \$3.00. The actual price charged by seed companies for hybrid seed is about \$70 per bushel.

In the case of F_1 hybrids, monopoly profits are possible because the farmer is no longer the principal competitor of his supplier, and the regular game of inter-enterprise competition is re-established. Market structure then determines the profitability of the seed industry. If entry into the market is easy and barriers to entry

low, competition and particularly price competition will be intense, with the result that margins and profits will be low. On the other hand, if a seed company achieves a monopolistic position, it will charge much more. Research and development play the same role as expenditure on sales promotion and, under modern economic conditions, are in reality key elements in the strategy by which enterprises build barriers to entry through this sales effort.

Not only do the marketing staff and sales personnel need new products or varieties to launch their sales campaigns, but the creation of new varieties depends on the past investment in research and development. A breeding program may take anything up to 10 years to yield a commercial variety. Seed maize companies devote 3 to 5 per cent of their total income to breeding activities. This may not seem a very impressive figure when compared to sales expenses (20 per cent or more of total income) but for any newcomer, it is a formidable obstacle: one would have to invest money for years, perhaps in vain, before getting any return on one's investment. Since the time horizon of entrepreneurs is short, oligopolistic market structures are very stable over time.

Because hybrids expand the physical market for seed to the total acreage sown and because they make it possible entirely to disconnect seed prices from actual costs of production, they open up enormous opportunities for profit for private enterprise. In the last 15 years, most of the vegetable crops have become hybrid, but the great breakthrough that private (and, in many cases, public breeders as well) are pursuing is the production of hybrid seed for wheat, barley and soya beans.

Insofar as the breeders' hope is to repeat the success achieved with maize, however, they are likely to be disappointed. Success, of course, means profits. The claims that hybrid maize increased yields by a large proportion are totally unsubstantiated. Moreover, it can be shown that the margin of the seed industry in the sale of a variety of seed depends on the multiplication rate (the ratio of the quantity harvested to the quantity sown) of the variety⁷. The higher the rate, the higher the potential profits. But wheat, barley and soya beans have multiplication rates which are one tenth or less of that of maize. This explains why the only other grain crops so far to have gone hybrid are sorghum, millet and sunflower, the multiplication rates of which are at least equal to that of maize. The multipli-

cation rate, the key to profits, has little to do with hybrid vigour, the biological phenomenon that allegedly justifies hybrid breeding schemes.

For wheat and barley, chemical inducing tillering is likely to increase the multiplication rate and therefore solve the seed companies' problem to some degree. If this is the case, the circle will be closed; the world seed industry has, in the past decade, passed into the control of powerful petrochemical and pharmaceutical corporations.

Conclusions

What we have described can be conveniently summarized in a set of concepts that provides a general framework for understanding some of the debates about biotechnology and property.

The specific factor of production implied in any biological process of production is genetic information. A wheat variety is a form of genetic information that makes it possible to make a better or more efficient use of the available resources, natural or technical. A strain of microorganisms used for the production of antibiotics is genetic information that makes it possible efficiently to use a substrate under precise conditions of pressure, temperature, acidity and so on.

Production in the economic sense involves a biological reproduction, which entails the production of genetic information in successive generations which can be identical, statistically similar or different to various degrees. For autogamous plants, barring mutations, the genetic information is identical. For a strain of microorganisms, mutations may occur during production, but the chance that they affect the useful part of the genome are negligible and, if they do, they can be selected against occasionally to keep their frequency low. For animals, the extent to which the useful genetic information is transmitted from parents to offspring depends upon the heritability of the character under consideration. In the case of openly-pollinated varieties of plants, both the gene frequency and random disorder are preserved; for production purposes, the offspring are similar to their parents only on the average.

Thus, except for hybrid varieties, production implies a diffusion of genetic information which is available to anybody and which becomes a public good. Street corner sellers of tomatoes, beans and potatoes transfer, *gratis*, the genetic information built into the commodity sold. The case of microorganisms is different; the commodity sold is not the living thing but

its product once it has been killed. Between the fermenters and the pill or injection, there may be several complicated steps; the customer is not, in any case, a direct competitor of his supplier. Nevertheless, the genetic information circulates easily; the folklore of the pharmaceutical industry is full of stories of visitors carefully wiping their noses and sending back the tissues to their own laboratories. The essential point is that the good concerned, genetic information, is difficult to control exclusively and tends to circulate freely.

Those who have been promoting breeders' rights in Europe, in the United States and in the rest of the world, particularly in the developing countries, as well as those who seek to patent other forms of life, have not paid attention to the innate contradiction in what they do.

Limiting the use of a good available in limitless quantities at no cost will not be socially useful, will limit the full use of biological potential only to what is patentable, will erect barriers to entry in branches of production where competition is necessary and will limit the free exchange of information between scientists so crucial to science. It would be unfortunate if the patenting system were surreptitiously extended to microorganisms. One may wonder what would have happened if the "Petition of Candle Merchants and Associated Industries against the Unfair Competition of the Sun", with its request that all openings of houses and buildings be covered, had been taken seriously. Frederic Bastiat's witty essay deriding the anti-free traders of his time applies exactly to the present situation.

Part of this work was financed by the Science-Technologie-Société Research Programme of CNRS. □

Jean-Pierre Berlan is at INRA-CEDEC, Université d'Aix-Marseille II, Marseille, France and Richard Lewontin is at the Museum of Comparative Zoology, Harvard University, Boston, USA.

1. *Actes des Conférences Internationales pour la Protection des Obtentions Végétales 1957-72*. (Procès-Verbal Résumé des Délibérations de la Première Session de la Conférence établie par le Secrétaire de la Conférence, M. Jauchviere, Paris 7-11 Mai 1975, page 24 (UPOV, Genève, 1974).
2. Barton, J. *Science* **216**, 1071-1075 (1982).
3. Testimony of representative of National Science Foundation in U.S. Congress, House of Representatives, Committee on Agriculture, Hearings before the Subcommittee on Department Investigations, Oversight, and Research, 96th Congress, First and Second Session, "Plant Variety Protection act amendment" (US Government Printing Office, 1980).
4. *Convention International pour la Protection des Obtentions Végétales*, Actes des Conférences Internationales, p.133-134, (UPOV, Genève, 1972).
5. UPOV, *Informations Générales*, Genève, 1975, p.7.
6. Barton, J. "The International Breeders' Rights System."
7. Berlan, J. P., *L'industrie des Semences*, Economie et Politique, *Economie Rurale*, 158: 18-28, 1983.

Fractal growth processes

Leonard M. Sander

Physics Department, University of Michigan, Ann Arbor, Michigan 48109, USA

The methods of fractal geometry allow the classification of non-equilibrium growth processes according to their scaling properties. This classification and computer simulations give insight into a great variety of complex structures.

ALMOST every theoretical tool of the condensed-matter scientist uses the assumption that the system considered is of high symmetry and is in equilibrium. These assumptions have led to enormous progress; however, to much, if not most, of the natural world such tools cannot be applied. Many systems that we would like to understand are very far indeed from perfectly ordered symmetry and are not even in local equilibrium. Perhaps the most extreme example is disorderly irreversible growth. We mean by this the sort of process which is very familiar in the formation of dust, soot, colloids, cell colonies and many other examples; roughly speaking, things often stick together and do not become unstuck. For example, a particle of soot grows by adding bits of carbon and coagulating with other particles in a random way. A possible result is shown in Fig. 1. We are thinking about cases which are, in some sense, as far from equilibrium as possible, and which have no obvious order.

It is remarkable that the introduction of simplified models has led to quite a good understanding of the morphology of such growth, despite the inapplicability of our usual modes of thinking. Here I will discuss this progress, drawing examples mostly from subjects which have traditionally interested physicists and chemists. However, disorderly growth is ubiquitous in the world around us, and is certainly not limited to inanimate matter. For example, some of the ideas which I will discuss, such as anomalous scaling in kinetic processes, will be useful to biologists. The purpose of the review is to introduce ideas from the area which may be of general use.

The key to our recent progress is the recognition that the most 'interesting' non-equilibrium structures (say, from a visual point of view) are not merely amorphous blobs; they still have a symmetry, despite their random growth habit, albeit a different one than they might have had, had they grown near equilibrium. For example, consider the soot of Fig. 1, or the electrolytic deposit of zinc shown in Fig. 2. Many people will be familiar with branched deposits such as this, and with similar looking objects which form on automobile windshields on cold mornings. In all these cases the structure is disordered, but it is not random. A manifestation of this is that each section of the picture contains holes in the structure comparable in size with that of the section itself. This can only occur if there are long-range correlations in the pattern; particles 'know' about each other over distances far in excess of the range of the forces between them. A truly random pattern, such as that of salt scattered on a table top, shows no such scaling of holes, and correlations are of short range only.

Studies of fractal growth have focused on two questions: how can we characterize and quantify the hidden order in complex patterns of this type, and when and how do such correlations arise? The answer to the first question is now relatively clear, and lies in an application of the fractal geometry of Mandelbrot¹. The next section gives a brief review of relevant aspects of this subject. The second question has received a partial answer in the formulation and analysis of models suitable for computer

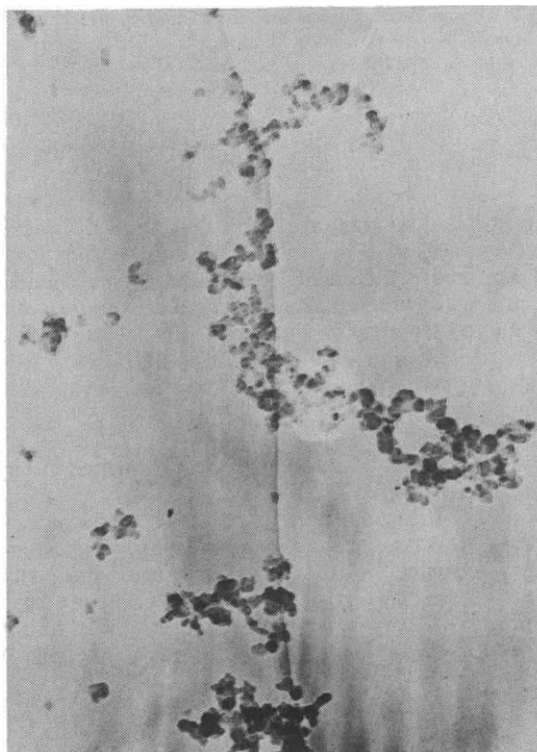


Fig. 1 Electron micrograph of soot. (Supplied by G. Smith, General Motors.)

simulation, which will also be reviewed. For more extensive treatments see refs 2-4.

Fractals and scale invariance

In pure mathematics, it has long been common to study certain 'pathological' geometric shapes that elude ordinary notions such as those of length and area. Figure 3 shows a famous example, which has, in some sense, infinite length, but zero area. It falls between our usual notions of line and solid. Mandelbrot¹ systematized and organized mathematical ideas concerning such objects due to Hausdorff, Besicovitch and others. But, more importantly, he pointed out that such patterns share a central property with complex natural objects such as trees, coastlines, patterns of stars and (as was later discovered) the non-equilibrium growths of Figs 1 and 2. This property is a symmetry which may be called scale invariance. These objects are invariant under a transformation which replaces a small part by a bigger part, that is, under a change in scale of the picture. Scale-invariant structures are called fractals.

There are a number of related properties which follow from the assumption of scale invariance. Consider, for example, the

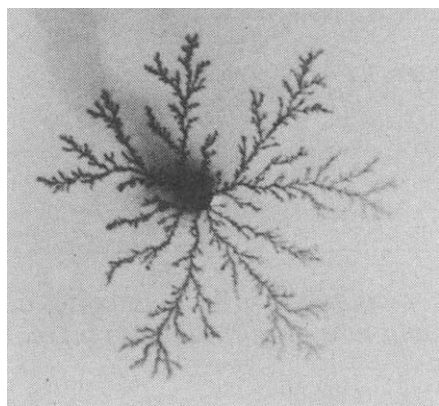


Fig. 2 Zinc electrodeposit produced in a thin cell under conditions of low ZnSO_4 concentration (0.01 mol l^{-1}). The outer electrode (not shown) is in the form of a ring 6.3 cm in radius. (Supplied by D. Grier, University of Michigan.)

density correlation function $c(r)$, of a fractal. This is defined as the average density of the object at distance r from a point on the object, and is a measure of the average environment of a particle. Clearly, $c(r)$ must reflect the scale invariance. It is easy to show that the only way that c may vary is as a power law in r ; any other function would have an intrinsic scale. It is convenient to write c in the following form:

$$c(r) = kr^{-(d-D)} \quad (1)$$

Here, k is a constant, and the exponent is written in terms of the dimension of space, d , and a new quantity, D , the fractal dimension. The reason for this terminology will become evident in a moment. As the objects we are dealing with are tenuous, $c(r)$ is a decreasing function of r : the average density decreases as the object becomes larger. Now consider how the total mass of the object, M , scales with the mean radius, R . We can estimate this by multiplying a typical density, from above, by the volume:

$$M(R) = KR^{D-d}R^d = KR^D \quad (2)$$

Here, K is another constant. We can now see why D is called a dimension. For an ordinary curve, $D=1$: twice the length gives twice the mass. For a disk, $D=2$. For simple objects D coincides with the usual notion of dimension. But in the cases we are discussing D is not an integer; it has been measured to be ~ 1.7 for the deposit in Fig. 2, and is 1.26 for the fractal of Fig. 3.

This anomalous scaling with radius, measured by D , is a very useful means of characterization because the fractal dimension is a 'robust' quantity. Like the famous scaling exponents of phase-transition physics, it has to do with long-range properties, indeed, with the relationship between properties at different scales. Thus we can expect it to be universal in the sense that it should be independent of the details of the interactions between the objects which stick together during the growth, of their detailed composition, and so forth. But, as we will see, the mechanism of growth does affect D .

Growth models

How might one visualize the growth of an object such as the electrolytic deposit in Fig. 2? As we are interested in long-range properties we can ignore the complications of electrochemistry and simply imagine that ions wander randomly in solution (in many cases the electric field is screened out so that this is a good approximation) and stick to the deposit when they happen to get near it.

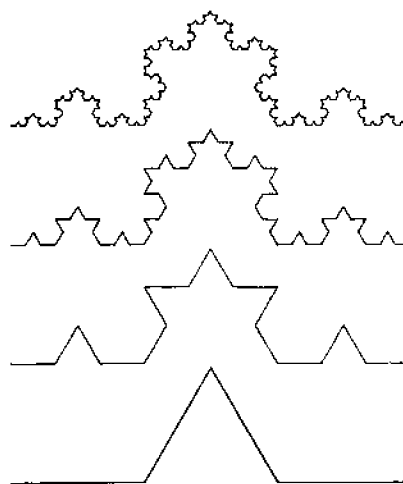


Fig. 3 Four stages in the growth of an exact fractal, the Koch curve. This and many other examples are discussed in ref. 1. The fractal dimension may be deduced by thinking of each picture as a part of the picture above, with a change of scale. For each scale change by three, we need four such parts. Thus, according to equation (2), $D = \log 4 / \log 3 = 1.26$.

To make a computer model which is a literal translation of this process we start with a centre. Then we liberate a diffusing particle, a 'random walker', and let it wander freely until it is within a fixed distance of the centre, where it sticks. Then we liberate another particle and let it walk until it sticks to the centre or the first particle, and so on. We may, for our purposes here, idealize the process of formation as being completely irreversible: we ignore the possibility that the particles rearrange after sticking to find a more energetically favourable location. This is the diffusion-limited aggregation (DLA) model of Witten and Sander^{5,6}. The application of DLA to electrodeposition is due to Brady and Ball⁷ and Matsushita *et al.*⁸.

Figure 4 shows the result of an extensive simulation according to the DLA rules; its resemblance to Fig. 2 is evident. Measurements of DLA clusters have shown them to scale according to the relations quoted above, with $D=1.7$ for $d=2$, and $D=2.4$ for $d=3$. Note that the structure is tenuous and open because holes are formed and not filled up. Filling up the holes would require wandering down one of the channels in the cluster without getting stuck on the sides; a random walker cannot do this.

There are several features of the DLA model which should be mentioned. Although it is simple to describe, no progress has been made towards 'solving' it. That is, although we suspect, on the basis of simulations, that DLA clusters are fractals, we cannot prove it. And we have no method of calculating D (or any other property): we must measure it. There are several reasons for this (I will mention a rather technical one below), the primary one being that DLA presents us with a situation in which our experience in equilibrium systems doesn't seem to help. Note that D , along with other scaling properties, arises in a non-trivial way from the kinetics of growth: there is no simple geometric argument with which to predict them.

The DLA model can be generalized in various ways, for example, to describe deposition on a surface⁹ rather than a point. A more profound generalization is to use the model to describe systems which apparently have nothing to do with particle aggregation, but which share the same universal properties. We may see how one is led to do this by observing^{5,10} that the probability, u , of finding a random walker at some point on its way to the aggregate has the following well-known properties: the flux of walkers, v , is proportional to the gradient of u , and, because walkers are absorbed only on the aggregate, this flux

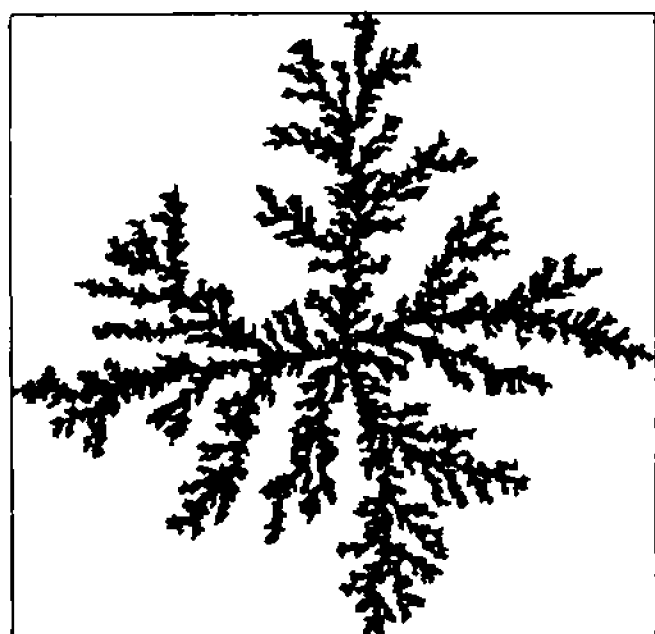


Fig. 4 A large DLA cluster ($\sim 50,000$ particles) grown on a square lattice. Note the resemblance to Fig. 2, and the beginning of distortion towards a dendritic outline, as discussed in the text. (Supplied by P. Meakin, Dupont.)

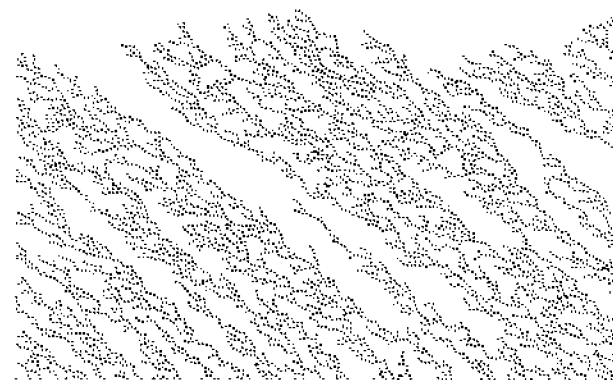


Fig. 5 Columnar microstructure in ballistic aggregation. Particles stick to the substrate and to each other after raining onto the structure in parallel trajectories at an angle to the vertical somewhat smaller than that of the columns. The fluctuations of the upper surface scale with the height for small height, and with the total width for large height. (Simulation performed by P. Ramanlal, University of Michigan.)

has no divergence:

$$v \propto \nabla u \quad (3)$$

$$\nabla \cdot v = \nabla^2 u = 0 \quad (4)$$

As walkers are not allowed to escape from the aggregate, we set $u = 0$ on the surface. The growth of the aggregate is given by the flux at its surface, that is, by ∇u .

As Niemeyer *et al.*¹⁰ pointed out, a set of equations of identical form govern dielectric breakdown of a solid if we ignore many short-range details. As we are looking for universal features, making such simple, indeed, crude approximations is justified. If we think of u as the electrostatic potential in a solid about to be destroyed by a discharge, its negative gradient is, of course, the electric field. But u then obeys the Laplace equation of electrostatics, which is of the same form as the steady-state diffusion equation, equation (4), above. The breakdown channel will grow in a way determined by the electric field, that is, the gradient of u , on its surface. If the growth rate is linear in the field, we expect to have exactly the same situation as in DLA, and indeed, direct solutions of the equations, as well as measurements of photographs of real discharges, give the same fractal dimension as DLA. Non-linear breakdowns (lightning in the atmosphere is probably an example) give rise to patterns with different values of D .

Paterson¹¹ noticed an even more remarkable manifestation of the wide applicability of the model. When a fluid flows under conditions of large friction, inertial effects are negligible and the flow rate can be taken to be proportional to the hydrostatic force, that is, to the gradient of the pressure: this is known as D'Arcy's law. The situation is commonly realized in the laboratory by letting fluid flow between thinly spaced plates, a so-called Hele-Shaw cell. In nature, the flow of crude oil through the porous rock in which it is found is an example of quite serious interest. Suppose we try to force such flow by blowing a bubble of air or another low-viscosity substance into the cell (or by pumping water into an oil field—a scheme known as enhanced recovery). It has long been known that the air will not uniformly displace the fluid; instead it will break up into a complex

structure with many arms¹², which are called 'viscous fingers'. This phenomenon has an obvious detrimental effect on enhanced recovery.

Paterson's¹¹ speculation was that the pattern of the viscous fingering would scale like DLA. His reasoning was as above: the pressure in an incompressible fluid obeys equation (4), with u now standing for pressure, because fluid, like particles, is conserved. D'Arcy's law is of the same form as equation (3). Once more, many details have been ignored. In particular, the role of surface tension in this and similar situations will be discussed below.

The reasoning has been verified most directly by Chen and Wilkinson¹³, who introduced discrete randomness into a Hele-Shaw cell—the effect should be that of the random arrivals of particles. Their patterns look almost exactly like Figs 2 and 4. Another experiment, by Nittmann *et al.*¹⁴, used the clever trick of eliminating surface effects by taking for the two fluids water and an aqueous polymer solution; the fluids are miscible but mix slowly. Once more the pattern of fingering resembled the simulations. There seems to be a source of randomness in this experiment, probably arising from the non-newtonian flow characteristics of the polymer solution; such shear thinning could amplify noise. Even more startling is the experiment of Ben-Jacob *et al.*¹⁵, who used a smooth Hele-Shaw cell, and one with a periodic pattern, with newtonian fluids. In some conditions they observed DLA-like scaling without an evident source of randomness, and without discrete 'particles'.

Experts will notice that equations (3) and (4) are of the same form (except for surface effects) as the description of solidification when the limiting factor in growth is diffusion of latent heat away from the surface of the growing crystallite. Why, then, does a snowflake (unlike the crystalline deposit of Fig. 2) not look like DLA, but is instead dominated by the crystal symmetry? I will return to this aspect of growth in the final section.

If particle aggregation doesn't need particles, what does it need? More generally, we can ask what different types of model give rise to scaling objects. For example, it is often the case that aggregates are formed by adding particles with a long mean free path, for example, in the formation of thin films by vapour-phase deposition¹⁶. In this case we may assume that the paths of the particles are straight lines. This model has become known as ballistic aggregation, and it has a number of very curious features. It is now known that the deposit itself is not a tenuous object but achieves a constant density^{17,18}. (In contrast, diffusion-limited growth on a surface⁹ yields an open deposit whose average density decreases with height.) It is of great interest to understand the upper surface of the film, which is a

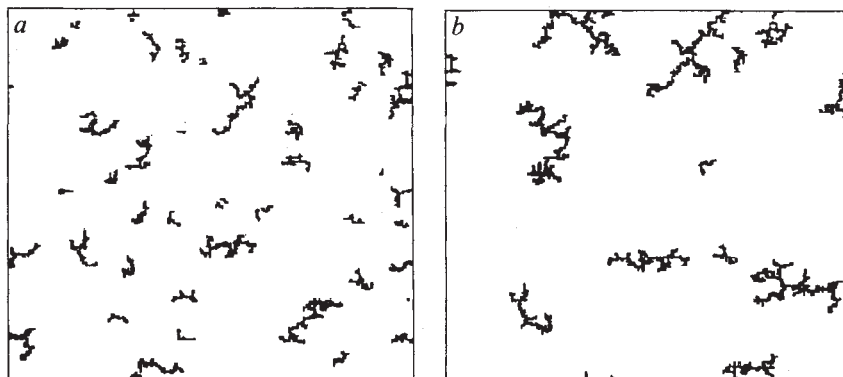


Fig. 6 Two stages in the formation of cluster-cluster aggregates: *a*, $t=3,669$; *b*, $t=17,409$. Note the resemblance of the clusters in *b* to the soot particle of Fig. 1. (Supplied by P. Meakin, Dupont.)

model of a random rough surface. It has been shown numerically¹⁹ that for normal incidence of the depositing particles this surface also has scaling properties: for example, the fluctuation of the height scales with a non-integral power of the height, for small height. This surface is probably not an ordinary fractal curve, like Fig. 2, but is probably an example of a self-affine fractal^{1,20}. 'Self-affine' means that the scaling in two different directions (width and height in this example) is different.

For non-normal incidence another effect appears, which is well-known in thin-film technology¹⁶. This is the columnar microstructure: the film spontaneously forms as a set of nearly parallel columns as it grows (see Fig. 5). The beginnings of a theory of this effect exist²¹, but it is not known what, if any, relationship these giant fluctuations have with the scaling fluctuations at normal incidence.

The simplest aggregation process of all was introduced into mathematical biology by Eden²². This is a model for the growth of a cell colony: a cluster is grown by adding particles at random to perimeter sites. Once again the object is compact, but the surface has interesting scaling properties which seem to be the same as for ballistic aggregates with normal incidence²³. Scaling is ubiquitous, and tends to have common features despite widely different details of growth.

We still have not described how soot forms. The structure of Fig. 1 is far more open than a DLA cluster: its fractal dimension is ~ 1.8 (DLA in three dimensions has $D=2.4$). Extensive measurements of soot²⁴, colloids²⁵ and other similar objects leads one to suspect that a different class of clusters is involved. In fact, we have omitted a central feature of the formation process of clusters which can coagulate, namely the aggregation of clusters with each other^{26,27}. Figure 6 shows two stages of a simulation of this process in two dimensions. We start with a vapour of freely moving particles which stick together whenever they come into contact, and then allow the clusters to continue to move with, perhaps, a smaller diffusion constant. The large fractals which are eventually formed have $D=1.4$. The corresponding simulations in three dimensions give $D=1.8$ and yield the open structure of real colloids and aerosols. At each stage of the process almost all of the clusters are of roughly the same size.

The open structure and low fractal dimension which characterize cluster-cluster aggregation are relatively easy to understand. It is difficult for a random-walking particle to penetrate a significant fraction of the radius of a growing cluster for particle aggregation; it is even more difficult for an aggregate of comparable size to do so. Thus, as aggregation proceeds, open, fluffy structures are produced.

One variant of this model which is worth mentioning is reaction-limited (chemically limited) aggregation²⁸. In many cases, because of the details of the growth process, the sticking is inefficient, and many attempts are required to form a new cluster. In the limit of a very large number of attempts, the fractal dimension increases from 1.8 to ~ 2 . Reaction-limited aggregation was probably discovered experimentally²⁹, before the simulations were done. Later experiments³⁰ have carefully

controlled the growth conditions and shown both growth mechanisms, and both types of geometry, in the same system for different growth rates. The encoding of kinetics in the scaling in a form independent of details should be a powerful tool for identifying growth mechanisms.

Attempts at theory

There is no general theory of irreversible growth. The descriptions given in the previous section must be regarded as a kind of phenomenology, albeit a useful one. We can point to situations in which there is scaling, but we are compelled to do experiments, either in the laboratory or on the computer to calculate anything. We do have a few analytical results, but they give only partial information.

The best understood type of aggregation is the cluster-cluster process. Suppose we assume, as stated above, that the dominant cluster-cluster collision is between clusters of similar mass. If we make the masses strictly equal we have a hierarchical model³¹. It is easy to believe then that we do have a fractal: agglomerating parts in this way is exactly how the artificial fractal of Fig. 2 was made. (Note that particle aggregation is not hierarchical, but it seems to be fractal nonetheless.) The specification of the size distribution of clusters in the vapour, and the verification of the hierarchical assumption, have been the objects of detailed studies³² which have shown that, indeed, the most common collision is between clusters of nearly equal mass. Some of these investigations use the techniques of colloid chemistry, in particular the Schmoluchowski kinetic equations, as well as computer simulations.

There remains the problem of finding the geometry of the clusters formed. Some progress has been made here because of a detail which allows one of the favourite tricks of the theoretical physicist to be applied. The real difficulty in visualizing the process is excluded volume, that is, the tendency of clusters to get in each other's way because they can attach to each other only on the outside. If they could attach anywhere it would be much simpler to sort out what is going on. It is quite obvious that excluded volume problems become less serious in high spatial dimensions: there are more ways into a three-dimensional cluster than into a two-dimensional one. Often, in equilibrium studies, it is found that for sufficiently large dimension of space, d , excluded volume is no problem at all: essentially any part of a cluster is accessible from outside. The dimension at which this starts to happen is called the upper critical dimension.

Above the upper critical dimension, calculations are simple, anomalous scaling is independent of d , and there exist methods (for equilibrium problems) which allow us to extrapolate to the physical world of $d=3$. For cluster-cluster processes, this is exactly what happens³³. The fractal dimension, D , for a cluster-cluster aggregate cannot grow above ~ 3.4 and it attains this value at about $d=7$. This is rather far from the real world, of course, and no one has yet figured out how to extrapolate.

The situation for particle aggregation is very different. Suppose that the entire cluster were to become accessible to added particles for a large enough value of d . Then the mass in the

interior would grow without adding to the volume. The cluster would quickly become so dense that it would no longer be accessible. Thus, there is no upper critical dimension for DLA. In fact, careful considerations of this sort can be turned into a bound³³ on D :

$$d-1 \leq D \leq d \quad (5)$$

The fractal dimension is never independent of the spatial dimension, and the standard technique cannot be applied.

Some progress has been made in the study of particle aggregation by exploiting the similarity of the process to the famous 'snowflake' problem, that is, the study of dendritic crystallization³⁴. We can see, for example, why tenuous structures are likely to arise in DLA and not in ballistic aggregation by noting that in the DLA case we have a growth instability of exactly the same form as the well-known Mullins-Sekerka³⁵ instability of crystal growth. The reasoning^{5,35} goes as follows: suppose we start with a smooth aggregate and ask why it grows sharp tips. If we start with a tiny bump on the surface it will be magnified into a tip by the fact that the bump will grow faster than the rest of the surface: it will catch random walkers more efficiently than the flat portions of the surface, and certainly much more efficiently than the holes in the aggregate. The analogous dielectric breakdown case will make this even clearer: recall that the growth rate of any point on the surface of the structure is proportional to the electric field there. Sharp tips have large electric fields (the lightning rod effect). They grow ever sharper and dominate the growth. In the viscous fingering problem the same instability arises because it is easy for viscous fluid to flow away from a growing tip. It is even possible to specify a relationship between D and the characteristic opening angle of the tips³⁶ by using the mathematical theory of lightning rods. Unfortunately, no one knows how to calculate these angles. In fact, recent work indicates that there is an array of sharp tips on the surface of the fractal DLA cluster whose distribution is itself fractal³⁷.

For ballistic aggregates or for the Eden model there is no growth instability: it is easy to see that a bump on the surface neither grows or shrinks, but just adds a uniform skin, and tips do not grow. The bulk of the material remains compact.

Fractals and snowflakes

In the last section we noted the usefulness of the analogy of DLA with the kind of solidification most familiar (at least to those in cold climates) in the formation of snowflakes, that is, branched (dendritic) crystals. But particle aggregates do not look like snowflakes. To be precise, in a typical dendrite, a growing tip forms by the Mullins-Sekerka instability but then stabilizes. It retains its shape and continues in a definite direction, although it may spawn side-branches as it grows. In DLA (and in, for example, the zinc deposit of Fig. 2) the tips repeatedly split and wander.

There are three obvious differences between DLA clusters and dendritic crystals: DLA has essentially zero surface tension, it has a significant source of noise in the discrete arrivals of the particles, and it has (at least in some versions of the model) no analogue of crystal anisotropy. Sorting out how these affect the process is a subject of current controversy and great intrinsic interest.

Surface effects can be added in various ways to DLA simulations^{5,38,39}; the result is to thicken the branches of the aggregate, but the scaling is unaffected for large sizes. Nor do surface effects, by themselves, make the equations of crystallization give rise to snowflakes^{40,41}. Instead, something unexpected happens: a growing tip with surface tension does not stabilize, but undergoes repeated splittings, which are caused by the surface tension itself. This is because surface tension slows the growth of sharply curved surfaces and the end of the tip is the most sharply curved. In order to make real dendrites, anisotropy arising from the crystal structure must be introduced. The relationship of anisotropy to tip-splitting was verified experimentally¹⁵ using fluid flow in a Hele-Shaw cell with a lattice of grooves.

How does this relate to DLA? It is common to do DLA simulations on a lattice (for convenience). Will the same thing happen here as in the noise-free case; that is, will stable tips form because of lattice anisotropy? It seems that the answer is yes^{42,43}: sufficiently large clusters on a lattice have the outline of a crystallite, with tip splitting only on a small scale. But why, without surface tension, do we ever get tip splitting? This is because noise due to the discreteness of the arriving particles can split the tips. This can be verified in various ways, for example, by experiments and calculations which vary the noise¹³ at fixed anisotropy. In cases where tip splitting is mainly due to surface tension rather than noise, will we get an object which scales? The answer to this question is not yet clear, but there are indications^{15,44} that there is scaling, and that it is close to that of DLA.

These considerations are of more than technical interest, because they show how small effects (such as anisotropy) can make qualitative changes in growth habit. A series of recent experiments^{45,46} have shown, for example, how changes in growth conditions, such as an increase in voltage in electrodeposition, can change a fractal pattern like Fig. 2 into an ordered dendritic crystal by increasing the effective anisotropy. This is a fascinating example of the competition between scaling symmetry and ordinary spatial symmetry.

This work was supported in part by NSF grant DMR 85-05474. I thank M. Sander and R. Merlin for comments on the manuscript, P. Meakin, G. Smith, D. Grier and P. Ramanlal for illustrations.

Note added in proof: There have been two interesting recent attempts^{47,48} to explicate the competition between anisotropy and noise.

- Mandelbrot, B. *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
- Family, F. & Landau, D. (eds) *Kinetics of Aggregation and Gelation* (North-Holland, New York, 1984).
- Pynn, R. & Skjeltorp, A. (eds) *Scaling Phenomena in Disordered Systems* (Plenum, New York, 1985).
- Stanley, H. & Ostrowsky, N. (eds) *On Growth and Form* (Nijhoff, The Hague, 1985).
- Witten, T. & Sander, L. *Phys. Rev. Lett.* **47**, 1400-1403 (1981).
- Witten, T. & Sander, L. *Phys. Rev. B* **27**, 5686-5697 (1983).
- Brady, R. & Ball, R. *Nature* **309**, 225-229 (1984).
- Matsushita, M., Sano, M., Hayakawa, Y., Honjo, H. & Sawada, Y. *Phys. Rev. Lett.* **53**, 286-289 (1984).
- Meakin, P. *Phys. Rev. A* **27**, 2616-2623 (1983).
- Niemeyer, L., Pietronero, L. & Wiesmann, H. *Phys. Rev. Lett.* **52**, 1033-1036 (1984).
- Paterson, L. *Phys. Rev. Lett.* **52**, 1621-1624 (1984).
- Saffman, P. & Taylor, G. *Proc. R. Soc. A* **245**, 312-329 (1958).
- Chen, J. & Wilkinson, D. *Phys. Rev. Lett.* **55**, 1892-1895 (1985).
- Nittman, J., Daccord, G. & Stanley, H. *Nature* **314**, 141-144 (1985).
- Ben-Jacob, E. *et al.* *Phys. Rev. Lett.* **55**, 1315-1318 (1985).
- Leamy, H., Gilmer, G. & Dirks, A. in *Current Topics in Materials Science* Vol 6 (ed Kaldis, E.) (North-Holland, New York, 1980).
- Bensimon, D., Shraiman, B. & Liang, S. *Phys. Lett.* **102A**, 238-240 (1984).
- Meakin, P. *Phys. Rev. B* **28**, 5221-5224 (1983).
- Family, F. & Viscek, T. *J. Phys. A* **18**, L75-L81 (1985).
- Voss, R., in *Scaling Phenomena in Disordered Systems* (eds Pynn, R. & Skjeltorp, A.) 1-11 (Plenum, New York, 1985).
- Ramanlal, P. & Sander, L. *Phys. Rev. Lett.* **54**, 1828-1831 (1985).
- Peters, H. P., Stauffer, D., Holters, H. P. & Loewenich, K. Z. *Phys. B34*, 399-408 (1979).

- Racz, Z. & Plischke, M. *Phys. Rev. A* **31**, 985-993 (1985).
- Richter, R., Sander, L. & Cheng, Z. *J. Coll. Interf. Sci.* **100**, 203-209 (1984).
- Weitz, D. & Oliveria, M. *Phys. Rev. Lett.* **52**, 1433-1436 (1984).
- Meakin, P. *Phys. Rev. Lett.* **51**, 1119-1122 (1983).
- Kolb, M., Botet, R. & Jullien, R. *Phys. Rev. Lett.* **51**, 1123-1126 (1983).
- Jullien, R. & Kolb, M. *J. Phys. A* **17**, L639-L643 (1984).
- Schaefer, D., Martin, J., Wiltrius, P. & Cannell, D. *Phys. Rev. Lett.* **52**, 2371-2374 (1984).
- Weitz, D., Huang, J., Lin, M. & Sung, J. *Phys. Rev. Lett.* **54**, 1416-1419 (1985).
- Botet, R., Jullien, R. & Kolb, M. *J. Phys. A* **17**, L75-L79 (1984).
- Botet, R. & Jullien, R. *J. Phys. A* **17**, 2517-2530 (1984).
- Ball, R. & Witten, T. *J. statist. Phys.* **36**, 863-866 (1984).
- Langer, J. *Rev. mod. Phys.* **52**, 1-28 (1980).
- Mullins, W. & Sekerka, R. *J. appl. Phys.* **34**, 323-329 (1963).
- Turkevich, L. & Scher, H. *Phys. Rev. Lett.* **55**, 1026-1031 (1985).
- Halsey, T., Meakin, P. & Procaccia, I. *Phys. Rev. Lett.* **56**, 854-857 (1986).
- Kadanoff, L. *J. statist. Phys.* **39**, 267-283 (1985).
- Viscek, T. *Phys. Rev. Lett.* **53**, 2281-2284 (1984).
- Brower, R., Kessler, D., Koplik, J. & Levine, H. *Phys. Rev. A* **29**, 1335-1342 (1984).
- Ben-Jacob, E., Goldenfeld, N., Langer, J. & Schon, G. *Phys. Rev. A* **29**, 330-340 (1984).
- Ball, R. & Brady, R. *J. Phys. A* **18**, L809-L813 (1985).
- Ball, R., Brady, R., Rossi, G. & Thompson, B. *Phys. Rev. Lett.* **55**, 1406-1409 (1985).
- Sander, L., Ramanlal, P. & Ben-Jacob, E. *Phys. Rev. A* **32**, 3160-3163 (1985).
- Grier, D., Ben-Jacob, E., Clarke, R. & Sander, L. *Phys. Rev. Lett.* **56**, 1264-1267 (1986).
- Sawada, Y., Dougherty, A. & Gollub, J. *Phys. Rev. Lett.* **56**, 1260-1263 (1986).
- Kertesz, J. & Vicsek, T. *J. Phys. A* **19**, L257-L262 (1986).
- Nittmann, J. & Stanley, H. E. *Nature* **321**, 663-668 (1986).

The Cretaceous/Tertiary boundary in the Gosau Basin, Austria

A. Preisinger*, E. Zobetz*, A. J. Gratz*[#], R. Lahodynsky[†], M. Becke[‡], H. J. Mauritsch[‡], G. Eder[§], F. Grass[§], F. Rögl^{||}, H. Stradner^{||} & R. Surenian[¶]

* Institut für Mineralogie, Kristallographie und Strukturchemie, Technische Universität Wien, Getreidemarkt 9, A-1060 Vienna, Austria

[†] Institut für Geologie, Universität Wien, A-1010 Vienna, Austria

[‡] Institut für Geophysik, Montanuniversität Leoben, A-8700 Leoben, Austria

[§] Atominstitut der Österreichischen Universitäten, A-1020 Vienna, Austria

^{||} Naturhistorisches Museum, Burgring 7, A1014 Vienna, Austria

[¶] Geologische Bundesanstalt, Rasumofskyg. 23, A-1031 Vienna, Austria

The Cretaceous/Tertiary boundary has been identified in the Gosau beds (Elendgraben) near Salzburg, Austria. The undisturbed 2-mm thick boundary clay in the palaeomagnetic G^- zone differs from the surrounding sediments in having significant colour, no biogenic calcite and different contents of rare-earth and siderophile elements, carbon and magnetic minerals. The clay also contains shocked quartz and plagioclase particles, and indicates a dramatic change in sedimentation caused by a short-lived event.

THE discovery of an iridium anomaly at the Cretaceous/Tertiary (K/T) boundary¹ has prompted a worldwide search for complete boundary sections. In Austria, prospective areas were examined for undisturbed boundary layers marking the event at the end of the Cretaceous period. To date, among several sections visited, only the Elendgraben in the Gosau Basin has been found to contain a complete sequence of sediments¹⁸.

We present here a brief but thorough study of the outstanding exposure at Gosau. The combination of geochemical, geomagnetic, mineralogical, palaeontological and sedimentological analyses provides a coherent and surprisingly precise picture of the events which took place at the K/T boundary. A more detailed presentation will appear elsewhere (A.P. *et al.*, in preparation).

Methods

In the course of detailed mapping of the Upper Gosau formation, hundreds of spot samples were collected and their fossil contents (nanoflora and microfauna) examined; this allowed the position of the boundary to be accurately located. In addition, 187 oriented rock cores were drilled for magnetostratigraphic studies. All core samples were thermally cleaned at 300 °C (and some cores at 360 °C) for rock magnetic analyses. The carrier minerals of the remanence are magnetite and/or titanomagnetite, as well as haematite. The mineral composition

was determined by means of X-ray diffraction, microprobe analysis, optical methods and electron microscopy. The samples were decalcified by decomposition with dilute acetic acid (10%), and the carbon contents determined by microanalyses.

Grain-size analyses of the sediments were performed with a sedigraph; chemical element analyses were carried out by X-ray fluorescence. For trace element analyses, neutron activation was performed at the Atominstitut der Österreichischen Universitäten, Vienna.

Site description

In the Elendgraben, near Gosau, which is a steep creek on the western slope of the Hornspitz-Höhbühel ridge (13°28'20" E; 47°33'23" N), flyschoid sediments of the Zwieselalm strata are exposed at an altitude of 1235 m (Fig. 1). The outcrop contains a >30-m thick sequence of southeastward-dipping sediments (dip angle 40°; dip azimuth 140°). The flyschoid sediments consist of alternating sandstones, siltstones, limestones, marly limestones and marls, which are partly turbiditic and partly hemipelagic. The sequence of these sediments indicates sedimentation in a deep-sea environment above the calcium carbonate compensation depth (CCD)². As shown by palaeocurrent patterns, the transported material, mostly metamorphic grains, was moving from south-east to north-west³.

The remarkably well preserved K/T boundary layers lie in a 16.7-m section of the palaeomagnetic reversed G^- zone (chron 29-R), 12.0 m above the magnetic polarity zone F_3^+ (top of chron

[#] Present address: Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA.

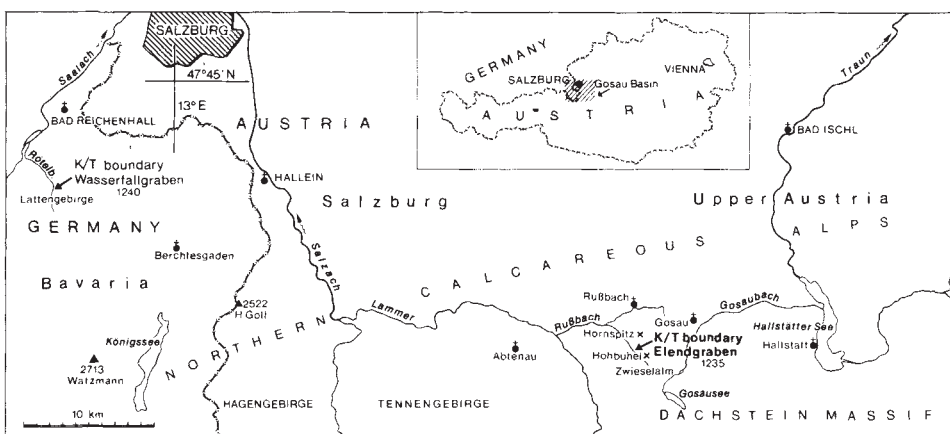


Fig. 1 Geographical position of the K/T boundary in the Elendgraben (Gosau Basin), south-east of Salzburg.

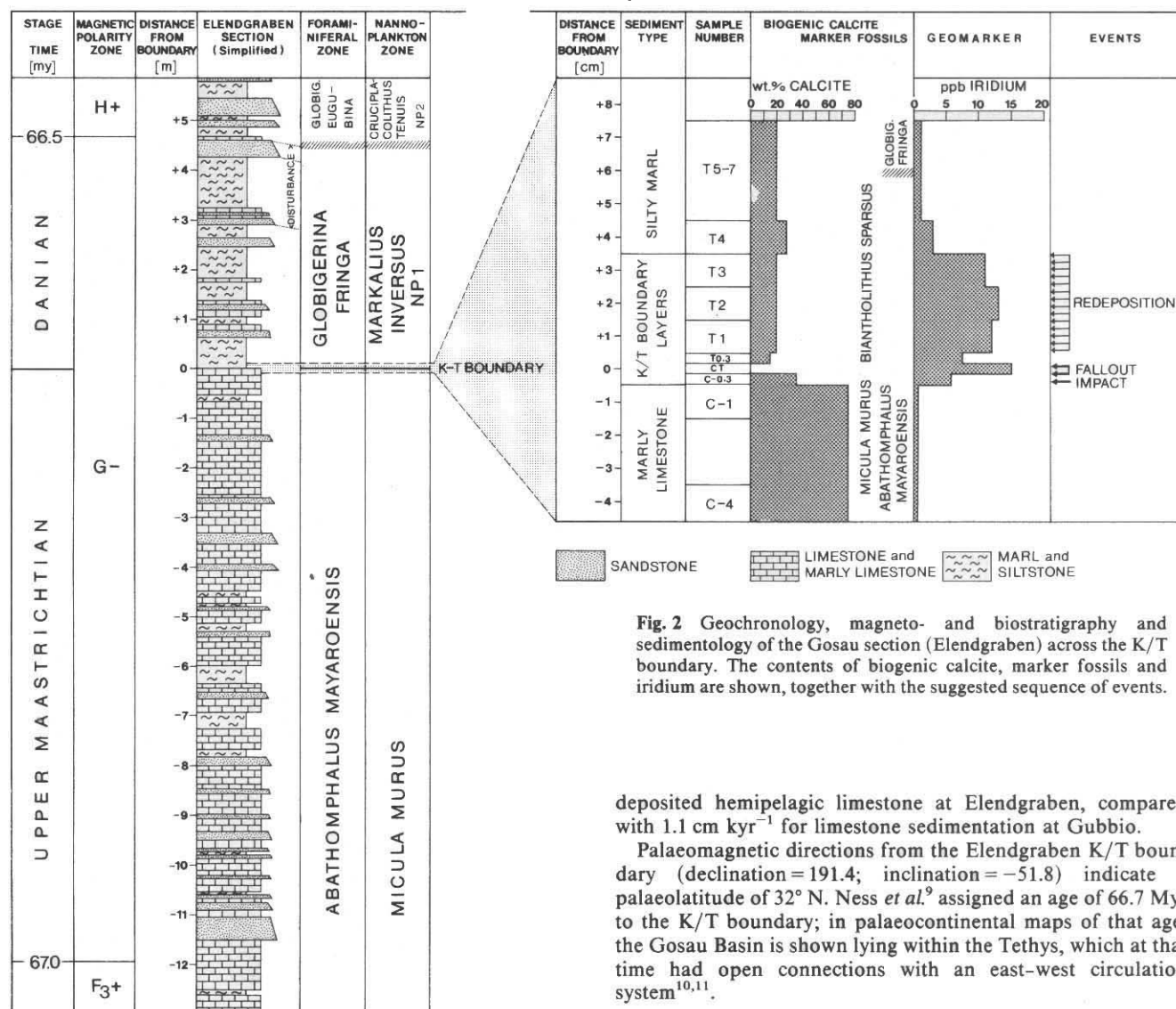


Fig. 2 Geochronology, magneto- and biostratigraphy and sedimentology of the Gosau section (Elendgraben) across the K/T boundary. The contents of biogenic calcite, marker fossils and iridium are shown, together with the suggested sequence of events.

deposited hemipelagic limestone at Elendgraben, compared with 1.1 cm kyr^{-1} for limestone sedimentation at Gubbio.

Palaeomagnetic directions from the Elendgraben K/T boundary (declination = 191.4° ; inclination = -51.8°) indicate a palaeolatitude of 32° N . Ness *et al.*⁹ assigned an age of 66.7 Myr to the K/T boundary; in palaeocontinental maps of that age, the Gosau Basin is shown lying within the Tethys, which at that time had open connections with an east-west circulation system^{10,11}.

Biostratigraphy

Two major fossil groups showing characteristic changes at the K/T boundary were studied in detail: the calcareous nanofossils and the planktonic foraminifera (Fig. 2). In the former group, the uppermost Maastrichtian assemblage contains the marker fossil *Micula murus*. In the marly limestone below the K/T boundary layers (samples C-1 to C-30), the nanofossils are partly recrystallized and poorly preserved; however, in the 3 mm of grey marl (C-0.3) of the K/T boundary layers underlying the light-brown K/T boundary clay (CT), they are well preserved. The K/T boundary clay itself is practically devoid of nanofossils, although some chamber fillings of foraminifera of uncertain origin are found. The foraminiferal microfauna below the K/T boundary layers belongs to the *Abathomphalus mayaroensis* zone. In the few centimetres above the K/T boundary clay, some reworked Maastrichtian foraminifera were found. Reworking of nanofossils from Cretaceous to Tertiary is also demonstrated by the presence of both Maastrichtian and Campanian coccoliths, such as *Aspidolithus parvus*.

Two centimetres above the K/T boundary clay, there is a sharp increase (by a factor of at least 10) in the frequency of the spherical calcareous shells of the dinoflagellate *Thoracosphaera operculata*, which is rather rare in the Upper Maastrichtian and becomes the dominant nannoplankton species in the K/T boundary layers. Perhaps the cells of *Thoracosphaera operculata* had better survival chances in case of a strong reduction of photosynthesis than many other unicellular species and also

30° N) and 4.7 m below the H^+ (base of chron 29-N), as shown in Fig. 2. Figure 3 shows a comparison of the sediment thickness of the G^- zone at Gosau (16.7 m) with those of other K/T boundary sites—such as Caravaca, Spain (16.8 m), DSDP site 524, South Atlantic (13.9 m), Stevns Klint, Denmark (8.8 m) and Gubbio, Italy (5.2 m)—which have approximately constant ratios of G^- (Tertiary) to G^- (Cretaceous)⁴⁻⁷. The overall thickness of the G^- zone depends on the type of sedimentation, specifically, the extent to which biogenic calcite is diluted by turbidites and/or continuously deposited terrigenous materials. The lowest sedimentation rates are found in pelagic deep-sea sediments, where the highest biogenic sedimentation rates are combined with the smallest amounts of terrigenous dilution (for example, the Gubbio G^- zone is only 5.2-m thick and contains 95% CaCO_3). The much greater thickness of the G^- zone at the Elendgraben section (Gosau) results from turbiditic and hemipelagic sedimentation. Discontinuously deposited turbidites account for ~48% of the thickness of these sediments (8.1 m); the remaining 8.6 m is taken up by hemipelagic limestone sediments containing 62% calcite (5.3 m biogenic calcite and 3.3 m terrigenous material); thus, the total amount of terrigenous material (bulk density $\approx 2.55 \text{ g cm}^{-3}$) deposited within the G^- zone is $8,400 \text{ kg m}^{-2}$.

The duration of the G^- zone is 470 kyr (ref. 8): this yields a mean sedimentation rate of 1.83 cm kyr^{-1} for the continuously

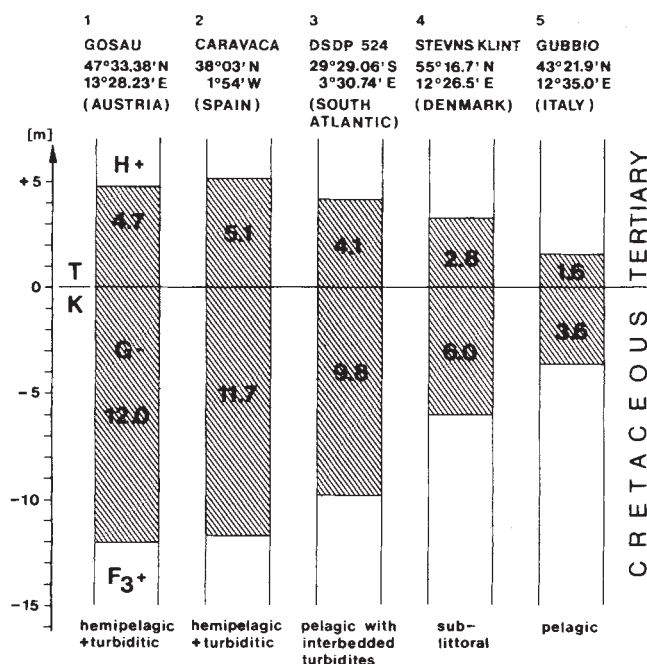


Fig. 3 Comparisons of the magnetic polarity zone G⁻ (chron 29-R) at different sites (Austria, Spain, South Atlantic, Denmark and Italy), with different types of marine deposits. The numbers above and below the K/T level indicate sediment thickness within that chron in Tertiary and Cretaceous, respectively.

had the ability to form close protective globular cysts. *Biantholithus sparsus*, the first Danian marker among the nannofossils (zone NP 1), is found immediately above the K/T boundary clay.

The first specimens of *Globigerina fringa* were discovered ~6 cm above the K/T boundary clay. In the Elendgraben outcrop, the lowest occurrences of the nannofossils *Cruciplacolithus primus*, *C. tenuis* (marking the base of the NP2 zone) and *Globigerina eugubina* lie very close to the change from the G⁻ to the H⁺ palaeomagnetic chron. There is strong agreement in biostratigraphical data between the Elendgraben section and equivalent sediments ~46 km away, in the Lattengebirge, Upper Bavaria¹².

Deep-sea sediments

The deep-sea sediments across the K/T boundary at the Elendgraben consist of marly limestone, K/T boundary layers and silty marl (Fig. 2), and do not contain turbidites.

The marly limestone (C-30 to C-1; thickness 30 cm) underlying the K/T boundary layers is consolidated and compact; it consists of 77% biogenic calcite (nannofossils and foraminifera), diluted by 23% non-biogenic components (quartz, feldspars consisting predominantly of oligoclase, mica, 14-Å-chlorite, smectite, illite and mixed-layer clays), the terrigenous sediments having the grain-size distribution: 3% sand, 30% silt and 67% clay. The Ir content is <0.5 p.p.b., the magnetic fraction is <0.02% and the organic carbon level is ~0.1%.

The K/T boundary layers (C-0.3 to T3) comprise ~4 cm of unconsolidated and porous material. The boundary layers comprise (from base to top) 3 mm of grey marl (C-0.3), 2 mm of light-brown K/T boundary clay (CT), 3 mm of grey marl (T0.3) and another 30 mm of grey marl (T1 to T3). The mean composition of layers T1 to T3 is 20% biogenic calcite, diluted by 80% non-biogenic components (quartz, feldspars, mica, 14-Å-chlorite, smectite, illite, mixed-layer clays and kaolinite) with mean grain-size fractions: 8% sand, 54% silt and 38% clay.

The 2 mm of light-brown boundary clay (CT) has a very different chemistry and mineralogy from the rest of the sediments

(see below); notably, it contains no biogenic calcite, and displays characteristic changes in the contents of certain trace elements ('geomarkers'), which reveal an event in Earth history by a marked increase or decrease in their quantity. Because of the worldwide presence of a significant iridium peak, the middle level of this thin layer (CT) is defined as the level of the Cretaceous/Tertiary (K/T) boundary. All distances given here are referred to this level.

The silty marl (T4 to T17) above the K/T boundary layers is 13 cm thick and friable. The content of biogenic calcite lies between 20 and 27%. The non-biogenic components have nearly the same mineral composition as the Cretaceous marly limestone but coarser grain sizes.

The grain size distributions show a mean size of quartz grains of ~8 µm for the Cretaceous marly limestone, and ≥20 µm for the Tertiary marls and silty marls.

Chemistry and mineralogy

We have used a combination of X-ray fluorescence, X-ray diffraction and microprobe analysis to determine the chemical contents and mineral compositions and abundances in decalcified samples. Some trace elements were analysed by neutron activation. Figure 4 shows the results of the chemical analyses of oxides and trace elements in samples C-4, C-1, C-0.3, CT, T0.3, T1, T2, T3, T4 and T5-7 across the K/T boundary, as well as information regarding their contents of transported quartz, expandable clays, kaolinite, magnetic fraction, elemental and organic carbon.

The samples of the Cretaceous marly limestone (C-4 and C-1) and the Tertiary silty marl (T5-7) are of similar composition. In contrast to these, however, is the dramatically different composition of the K/T boundary layers (C-0.3 to T3), especially the boundary clay (CT) with its extraordinary chemical and mineral composition. The 2-mm boundary clay (CT) contains no biogenic calcite (Fig. 2); a minimum of quartz, plagioclase, mica and 14-Å-chlorite; a minimum of SiO₂, Na₂O and K₂O; relatively high amounts of kaolinite and expandable clays with high water content (Ca-smectite and mixed-layer clays); and peak values of magnetic minerals, such as exsolved ilmenite (FeTiO₃) in magnetite (Fe₃O₄), formerly titanomagnetite. The weight percentages of the authigenically formed minerals (kaolinite, Ca-smectite and mixed-layer clays) total ~94%, which, like the Ir content of 14.5 p.p.b. (parts per 10⁹), is a peak value. The high Ir content of the boundary clay (CT) was shown to be related to the titanium mineral and not to the magnetite. A marked increase of carbon, as determined by microanalyses, is also apparent in the K/T boundary layers.

The bulk of the 2-mm-thick boundary clay (CT) is comprised of fallout material (mainly glass and its derivatives, and shocked and deformed minerals). The succeeding sequence of K/T boundary layers (T0.3 to T3) also contains reworked, transported and redeposited material, indicating different sedimentation processes.

The >20-µm fractions of decalcified marl and clay (C-0.3 and CT) were examined for the presence of shocked particles (Fig. 5), using a polarizing microscope equipped with a universal stage, and also by means of X-ray diffraction and transmission electron microscopy (TEM). X-ray analysis shows that 70% of the fraction is quartz, the remainder consisting of feldspar and kaolinite. Whereas the quartz in the underlying Maastrichtian (C-1 and C-4) is a fine groundmass with <10% of 10-µm clasts (undeformed to slightly deformed), the quartz of the samples C-0.3 and CT is clearly shocked. Shocked quartz was first reported at the K/T boundary in Montana¹³.

In the shocked quartz of the Elendgraben, shock features include lamellae indexed to {1013}, {1011}, {1012}, {0001}, {1122} and {1121}, planar cleavages shown to be microfaults in laboratory shock specimens, symplectic material, and widespread vitrification¹⁴. Eighty-two of more than 100 grains in sample C-0.3 were identified as quartz, showing the following

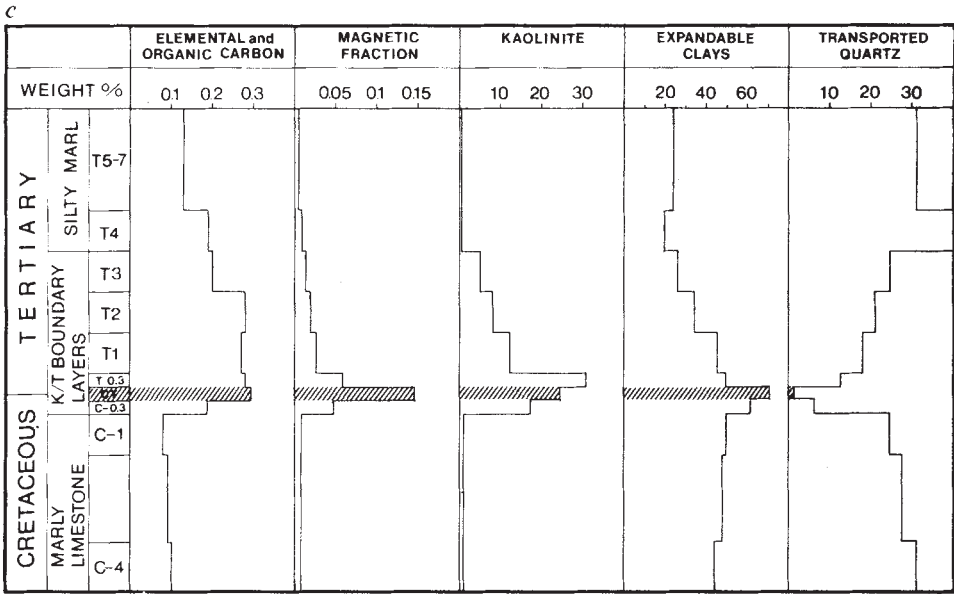
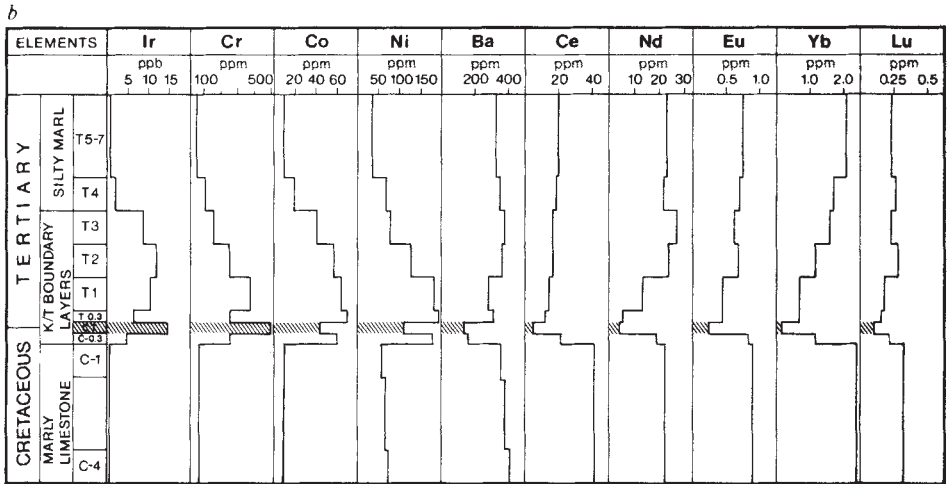
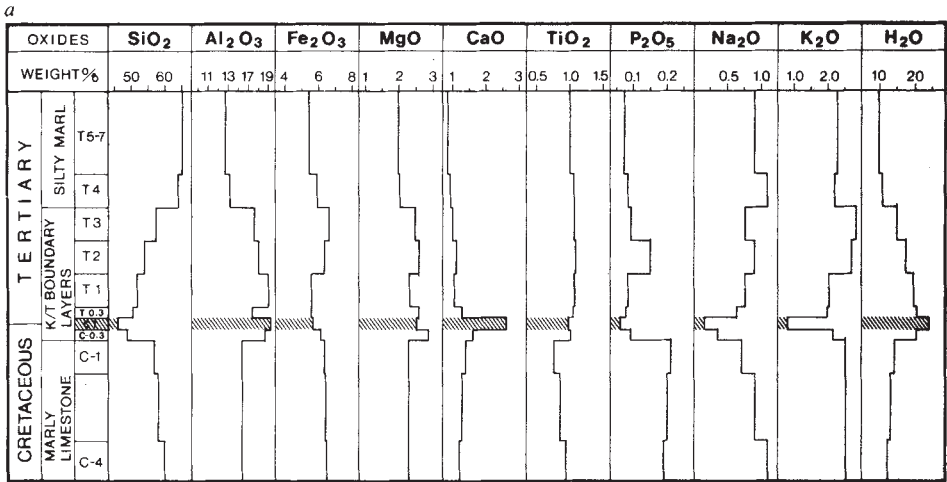


Fig. 4 Analyses of samples across the K/T boundary (Elendgraben), after decalcification. a, Distribution of major-element oxides; b, contents of trace elements (geomarkers); c, biogenic and non-biogenic material. Shaded area, K/T boundary clay (CT).

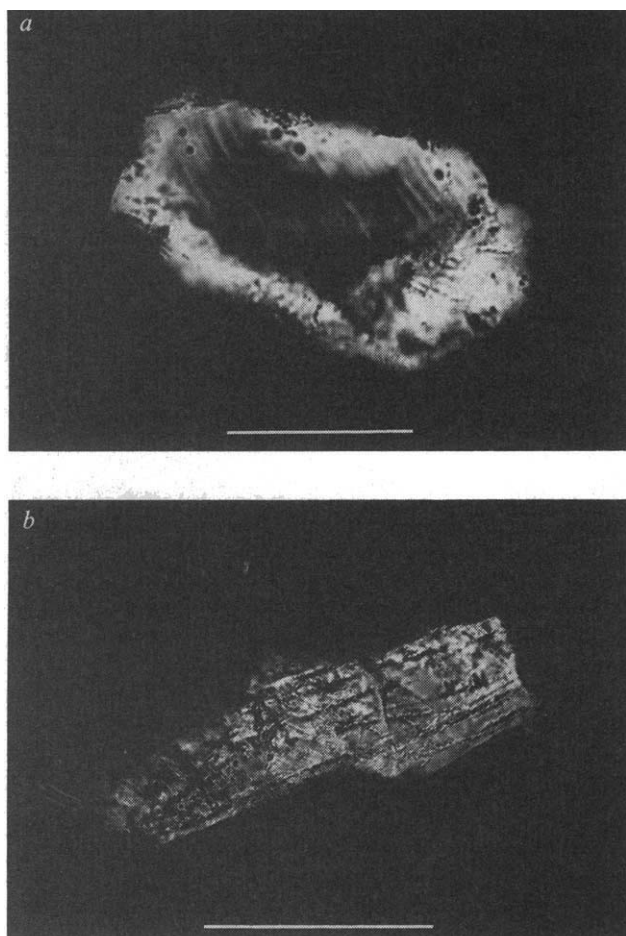
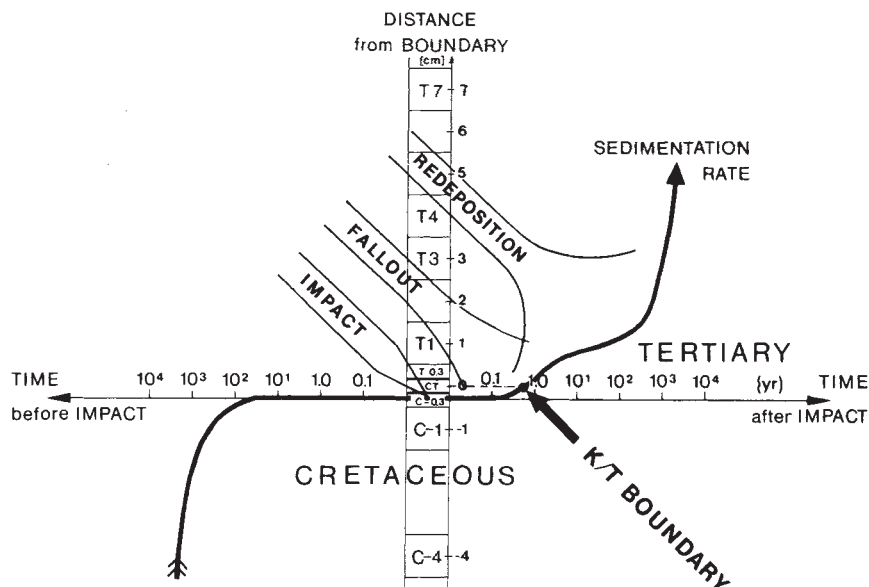


Fig. 5 Photomicrographs of minerals deformed by the K/T boundary event (Elendgraben). *a*, Shocked quartz grain (sample C-0.3) with lamellae; *b*, shocked plagioclase grain (sample C-0.3), with numerous planar features. Scale bars, 10 μ m.

features: 8 grains contain shock lamellae; 40 grains contain microfault sets; 26 grains contain isolated microfaults, decorated (partially annealed) microfaults, and/or planar edges; 3 grains are largely vitrified, symplectic, or melted; 3 grains show only internal fracturing; and 2 grains show no distinct signs of shock. All grains show mild to extreme undulatory extinction. These results are confirmed by TEM, which showed an enormous quantity of amorphous material; microcrystallites distributed in amorphous material; consistently planar grain edges, usually decorated with glass; planar glassy areas; and no or few dislocations. It appears that essentially all quartz grains examined are shocked and broken along shock microfaults. Furthermore, material too fine or too coarse for the survey shows a similar distribution of effects, so that the majority of 20- μ m quartz—that is, approximately half of the 11% quartz from sample C-0.3—is shocked to varying degrees (Fig. 5a).

Feldspar grains show similar effects (Fig. 5b), including planar features and edges, undulatory extinction, and widespread glass in TEM. Phyllosilicates, however, have none of the distinctive shock kink bands. Material from the boundary clay (CT) is similarly shocked, but contains conglomerates of shocked grains, more rounded grains, and recrystallized material with relict microfaults and lamellae. This material was shocked and then annealed and deposited. Also, large, euhedral, twinned calcite crystals are found.

The presence of shocked particles in sediments, in connection with the mass extinction of foraminifera and nannoplankton is considered to be evidence for an extraordinary event. Furthermore, the enrichment of geomarkers such as Ir, Cr, Co and Ni, and the depletion of geomarkers such as Ba and the rare-earth elements (Fig. 4), indicate that the event marking the K/T boundary could be caused by an impact of an asteroid of nearly chondritic composition, as suggested previously^{15,16}. The mass of such an asteroid can be estimated from the total amount of excess Ir deposited in the K/T boundary layers. Such a calculation¹⁷ yields a diameter of >8 km, which is in agreement with data in ref. 1. The shocked quartz and plagioclase, the rock fragments, the Ca-smectite, the twinned calcite crystals, the excess of Mg and Ca ions in the K/T boundary layers, and the global distribution of a large amount of chondritic material indicate that the event was initiated in the ocean. The amount of the crater material that was ejected into the atmosphere after the impact and distributed across the globe can be calculated from the fallout and its derivatives (kaolinite, mixed-layer clays and Ca-smectite), plus the amounts of reworked and redeposited



material¹⁷. This amount deposited as a compact sediment layer across the globe corresponds to a thickness of ~9 mm.

Conclusions

The distribution of the geomarkers and the mineral components (Figs 2, 4) in the K/T boundary layers (C-0.3 to T3) suggests that there was a short period of 'fallout', followed by a redeposition of the fallout material over a longer period of time. On the basis of the analyses given here, Eder and Preisinger¹⁷ have concluded that the time required for the deposition of the K/T boundary layers was much shorter than for any other layers in the sequence. As shown in Fig. 6, where sediment thickness is plotted on a linear and time on a logarithmic scale, the impact itself ($t=0$) occurred within layer C-0.3 and is not identical with the K/T boundary indicated within layer CT. This is consistent with the idea that the material ejected into the atmos-

phere returned to the Earth over a period of time, which would explain why the Ir peak is not synchronous with the impact.

The combination of the analytical results and the time-structure analysis of the K/T event suggest that, 66.7 Myr ago, an impact of an asteroid took place, producing a worldwide geochemical and mineralogical anomaly within a very short stretch of time and causing conditions for mass extinctions of large quantities of biota.

We thank Doz. Dr G. Kurat and Dr F Brandstätter (Museum for Natural History, Wien) for microprobe analyses, Dr P. Pongratz (Institute for Applied Physics of the Technical University, Wien) for transmission electron microscopy, Dr J. Zak (Microanalytical Laboratory of the University, Wien) for carbon microanalyses, the Research Center Seibersdorf, Austria, for radiation time and the Fonds zur Förderung der wissenschaftlichen Forschung in Österreich (FWF) for financing part of the nannoplankton research (project no. 2659).

Received 10 February; accepted 19 June 1986.

1. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095-1108 (1980).
2. Faupl, P. *Mitt. Ges. geol. Bergh. Öst.* **25**, 81-110 (1978).
3. Lahodinsky, R. *Verh. geol. Bundesanst., Wien*, 109-110 (1979).
4. Smit, J. *Spec. Pap. geol. Soc. Am.* **190**, 329-352 (1982).
5. Hsü, K. J., McKenzie, J. A. & He, Q. H. *Spec. Pap. geol. Soc. Am.* **190**, 317-328 (1982).
6. Mörmner, N.-A. *J. Geol.* **90**, 564-573 (1982).
7. Roggenthen, W. M. & Napoleone, G. *Geol. Soc. Am. Bull.* **88**, 378-382 (1977).
8. La Breque, J. L., Kent, D. V. & Cande, S. C. *Geology* **5**, 330-335 (1977).
9. Ness, G., Levi, S. & Couch, R. *Rev. Geophys.* **18**, 753-770 (1980).
10. Smith, A. G., Hurley, A. H. & Briden, J. C. in *Phanerozoic Palaeogeographic World Maps* (Cambridge University Press, 1981).
11. Haq, B. U. *Oceanol. Acta* **71-82** (1981).
12. Herm, D., von Hillebrandt, A. & Perch-Nielsen, K. *Geologica bav.* **82**, 319-344 (1981).
13. Bohor, B. F., Foord, E. E., Modreski, P. J. & Triplehorn, D. M. *Lunar planet. Sci.* **15**, 74-75 (1984).
14. Gratz, A. J. *J. Noncryst. Solids* **67**, 543-558 (1984).
15. Smit, J. & Ten Kate, W. G. H. Z. *Cretac. Res.* **3**, 307-332 (1982).
16. Hsü, K. J. *Patterns of Change in Earth Evolution*, 63-74 (Springer, Berlin, 1984).
17. Eder, G. & Preisinger, A. *Naturwissenschaften* (in the press).
18. Stradner, H. et al. *Terra Cognita* **5**, 247 (1985).

Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria

Anthony G. DiLella*, Joshua Marvit*, Alan S. Lidsky*, Flemming Güttler* & Savio L. C. Woo*

* Howard Hughes Medical Institute, Department of Cell Biology and Institute of Molecular Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

† John F. Kennedy Institutiet, Gl. Landevej 7, DK-2600 Glostrup, Denmark

The first phenylketonuria mutation identified in the human phenylalanine hydroxylase gene is a single base substitution (GT→AT) in the canonical 5'-splice donor site of intron 12. Direct hybridization analysis using specific oligonucleotide probes demonstrates that the mutation is tightly associated with a specific restriction fragment-length polymorphism haplotype among mutant alleles. The splicing mutation is the most prevalent phenylketonuria allele among Caucasians, and the results suggest the possibility of detecting carriers of the genetic trait who have no family history of phenylketonuria.

CLASSICAL phenylketonuria (PKU) is an autosomal recessive human genetic disorder caused by a deficiency of hepatic phenylalanine hydroxylase (PAH, phenylalanine 4-monooxygenase, EC1.14.16.1). In the normal human liver, PAH catalyses the rate-limiting step in the hydroxylation of phenylalanine to tyrosine. The reduction in PAH activity causes an accumulation of serum phenylalanine, resulting in hyperphenylalaninaemia and abnormalities in the metabolism of many compounds derived from the aromatic amino acid. PKU is the most common inborn error in amino-acid metabolism. Its incidence among Caucasians ranges from 1:4,500 in Ireland to 1:16,000 in Switzerland, with an average incidence of about 1:8,000 in the United States. The mutant gene frequency is such that 1 in 50 Caucasians is a carrier of the disease trait. Without early detection of PKU, followed by rigid implementation of a restricted diet low in phenylalanine during the first decade of life, affected

children develop severe mental retardation (for reviews see ref. 1-3).

Although neonatal screening for PKU is routinely carried out in Western countries, there has been no conventional means of identifying affected fetuses. We used a full-length human PAH complementary DNA clone⁴ to identify and map eight restriction fragment-length polymorphisms (RFLPs) at the human PAH locus⁵⁻⁷. These RFLPs segregate in a mendelian manner and concordantly with the mutant alleles in PKU kindreds. The frequency of the RFLPs in the PAH gene is such that the observed heterozygosity in the general population is about 87% (ref. 8). Thus, prenatal diagnosis can be performed in most PKU families by RFLP analysis⁹. However, the analysis is limited in that it can only be used in families with a history of PKU.

We recently used RFLPs to identify 12 haplotypes of normal and PKU alleles in the Danish population, and observed a

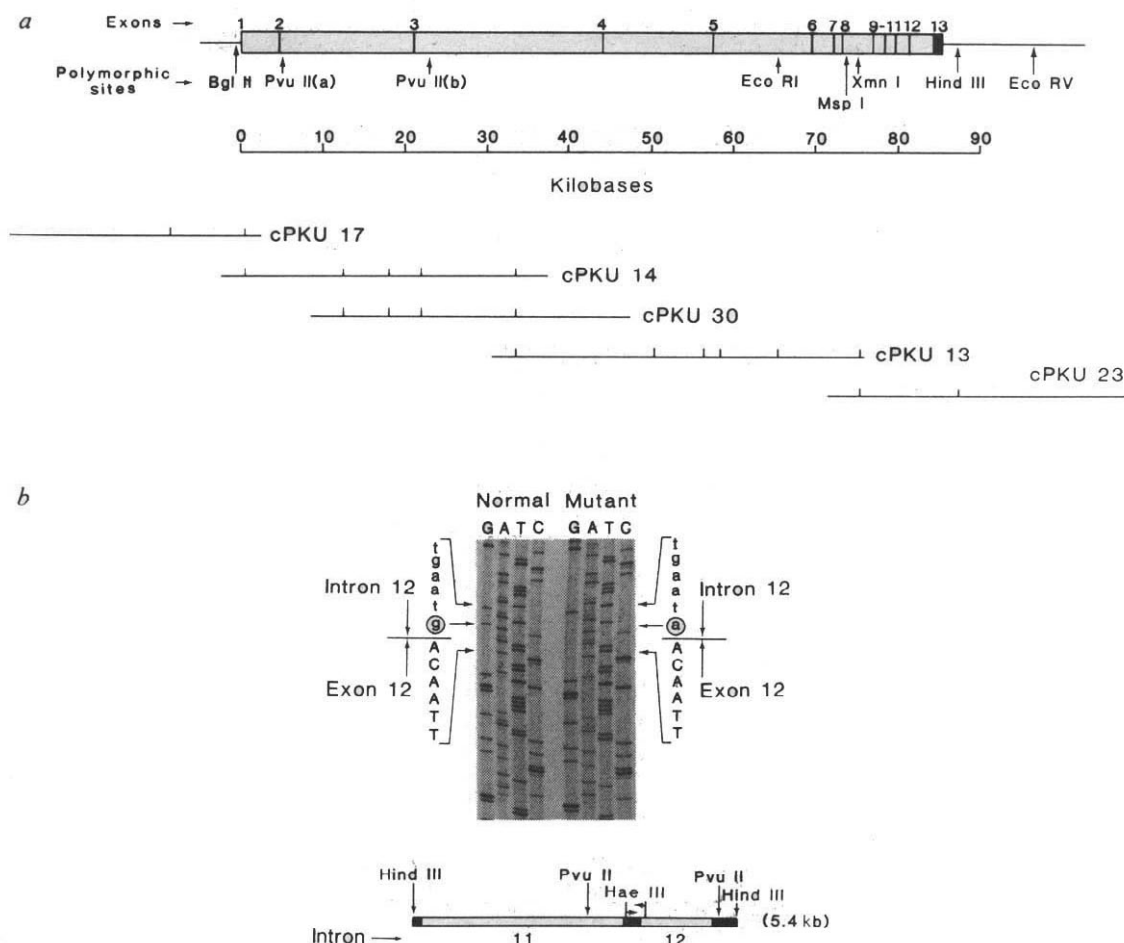


Fig. 1 Identification of a PKU splicing mutation. **a**, Physical map of the mutant gene. The maps of five overlapping cosmid clones (cPKU 17, cPKU 14, cPKU 30, cPKU 13 and cPKU 23) were used to characterize the mutant gene. The subcloned *Eco*RI fragments containing exon sequences were used for Sanger dideoxynucleotide sequence analysis¹⁸ as described previously for the normal PAH gene⁷. The detailed structure of the mutant gene in the 5' to 3' orientation and the location of RFLP sites⁷ are also shown. The cosmid library construction and gene characterization were carried out as described previously^{7,20}. **b**, Sequence analysis of the 5'-donor splice site of exon 12. The 114-base-pair (bp) *Hae*III fragment (bottom) containing 90 bp of exon 12 and 24 bp of the downstream intron was inserted into the *Sma*I site of M13mp18¹⁹ and sequenced in both directions (arrows) by the dideoxynucleotide chain-termination method¹⁸. A comparison of the normal and mutant sequence at the intron/exon border illustrates the single base substitution.

strong association among distinct RFLP haplotypes and PKU alleles⁸. A close association has been observed previously between RFLP haplotypes of the β -globin locus and specific β -thalassaemia mutations in different ethnic populations (for review see ref. 10). Whether such tight association between RFLP haplotypes and specific mutations exists in PKU due to linkage disequilibrium can be verified only by molecular analysis of the mutant genes. Having established the molecular structure and RFLP haplotype linkage map of the PAH gene⁷, this important issue can be addressed by cloning and characterizing the mutant PAH alleles of the prevalent RFLP haplotypes.

Isolation and sequencing of a mutant PAH gene

RFLP haplotypes for both PKU and normal PAH genes have been determined for 33 informative Danish PKU families⁸. About 75% of the normal and 90% of the mutant PAH genes are confined to four common RFLP haplotypes (Table 1; haplotypes 1, 2, 3 and 4). Thirty-eight percent of the mutant alleles have a single haplotype (haplotype 3) that is relatively rare among the normal gene pool (3%). A cosmid genomic

DNA library was thus constructed from leukocyte DNA isolated from a PKU individual who is homozygous for haplotype 3. The library was screened with a full-length human PAH cDNA probe⁴ and several corresponding genomic clones were isolated (Fig. 1a). To identify the mutation causing PKU, exon-containing regions of the gene were subcloned into the vector M13mp18 for sequence analysis. The mutant gene sequence is identical to that of the normal PAH gene reported previously⁷ except for a silent nucleotide substitution (A $\rightarrow$ G) in the third base of codon 232 (Gln)^{4,7}, and a G to A transition at the 5' splice donor site of intron 12, altering the obligatory G-T donor dinucleotide to A-T (Fig. 1b). Gene transfer and expression experiments demonstrated that this mutation results in abnormal processing of PAH messenger RNA and loss of PAH activity (J.M. *et al.*, unpublished results).

Design and testing of probes

Figure 2b shows the nucleotide sequence at the mutation site in the gene; this did not result in any alteration in restriction recognition sequence from the normal PAH gene. Since a single base-pair mismatch is sufficient to destabilize duplex structures

Table 1 RFLP haplotype distribution of the PAH alleles in Denmark

Haplotypes	Normal alleles		Mutant alleles	
	Number*	%	Number*	%
1	23	34.8	12	18.2
2	3	4.6	13	19.7
3	2	3.0	25	37.9
4	21	31.8	9	13.6
5	7	10.6	0	0
6	0	0	2	3.0
7	7	10.6	1	1.5
8	1	1.5	0	0
9	0	0	1	1.5
10	1	1.5	0	0
11	1	1.5	1	1.5
12	0	0	2	3.0

* Number of alleles observed in a total of 66 chromosomes analysed.

All tested families were Caucasian and had no history of consanguinity. Clinical criteria of the probands as affected PKU individuals and parents as heterozygotes in the families have been reported previously¹⁵.

in DNA hybrids^{11,12}, oligonucleotide probes can be used to distinguish between normal and mutant alleles in the human genome. Thus, we synthesized a normal-specific probe (21-mer) which is complementary to the sense-strand of the normal gene (Fig. 2b) and forms a C-A mismatch at the mutation site with the mutant gene. For the mutant-specific probe, a 21-mer of the sense-strand sequence was synthesized. At the mutation site this probe forms a C-A mismatch with the normal gene sequence (Fig. 2b).

The fidelity of the two oligonucleotide probes in detecting the respective sequences was then tested using the normal and mutant genomic DNA clones. Due to the high A+T contents of the normal and mutant probes (76% and 81%, respectively), we used the base-composition-independent oligonucleotide hybridization method^{13,14} to analyse the point mutation. *PvuII* digestion of the 12-kilobase (kb) *EcoRI* fragment isolated from both the normal (N) and mutant (M) gene clones generated a 2-kb fragment containing the entire exon 12 plus flanking intronic sequences. Under the hybridization conditions used, the normal probe hybridized only to the 2-kb *PvuII* fragment of the normal gene (Fig. 2a, middle panel), and the mutant probe hybridized specifically to the mutant gene (Fig. 2a, right panel).

Mendelian segregation of PAH alleles

Using the appropriate hybridization conditions described in Fig. 2 legend, the synthetic oligonucleotide probes were used to identify the mutant alleles in genomic DNA and to analyse their segregation in PKU kindreds. The first family analysed was the one from which the mutant allele had been isolated and characterized. In this family, both parents contain a mutant haplotype 3 gene which hybridized to the mutant probe (Fig. 3a, lanes 1, 2), suggesting that both mutant alleles in this family may contain the same mutation. Since this disorder is autosomal recessive in nature, both parents are obligate carriers of the PKU trait and both must also contain a normal gene. In this case, both normal alleles in the two parents corresponded to haplotype 4, and both hybridized to the normal probe as expected (Fig. 3a, lanes 7, 8). There are two affected individuals in this family; both are homozygous for the mutant haplotype 3 alleles which hybridized strongly to the mutant probe (Fig. 3a, lanes 3, 4), but not to the normal probe (Fig. 3a, lanes 9, 10).

The results confirm that both mutant alleles in this family contain the same mutation. An unaffected sibling in this family is homozygous for the normal haplotype 4 alleles which hybridized to the normal probe (Fig. 3a, lane 11) but not to the mutant

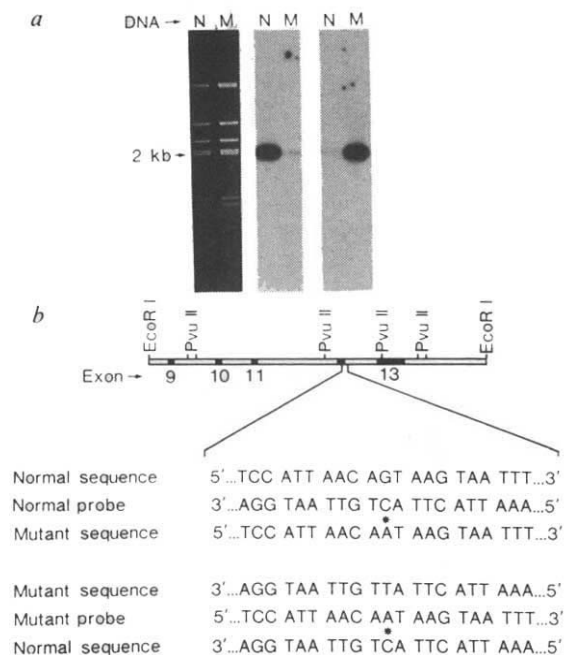


Fig. 2 Oligonucleotide specificity for the detection of the PKU mutation in cloned DNA by direct-gel hybridization. **a**, The 12-kb *EcoRI* fragment containing exons 9 to 13 was isolated from normal (N) (ref. 7) and mutant (M) cosmid clones and inserted into plasmid pBR322. *PvuII* plus *EcoRI* digests of the recombinant clones were electrophoresed on 1% agarose gels, stained with ethidium bromide (left panel), and processed for direct-gel hybridization as described previously¹¹. Dried gels were hybridized to oligonucleotide probes specific for normal (middle panel) and mutant (right panel) cloned PAH genomic DNA sequences. Exon 12 plus flanking intronic sequences are contained in a 2-kb *PvuII* fragment. The gels were overexposed intentionally in order to demonstrate the differential hybridization signals generated by the two oligonucleotide probes. **b**, Oligonucleotide probes used to detect the normal and mutant 5'-donor splice site of exon 12. The asterisks denote the C-A mismatch between the normal probe and mutant gene sequence, and the mutant probe and normal gene sequence.

Methods. Oligonucleotides were synthesized by an automated (SYSTEC, Inc.) solid-phase phosphite triester method. High specific activity probes (10^{10} c.p.m. μg^{-1}) were generated by the primer extension method¹². The normal probe was synthesized on a 21-base template (5'-TCCATTAACAGTAAGTAATTT-3') by hybridized 9-base primer (3'-TTCATTAATA-5'), while the mutant probe was synthesized by extension of a 9-base primer (5'-TCCATTAATA-3') hybridized to a 21-base template (3'-AGGTAATTGTTATTCATTAATA-5'). Dried gels were hybridized overnight at 37°C in 6×NET (0.9 M NaCl, 6 mM EDTA, 0.5% SDS and 0.09 M Tris, pH 7.5) containing 0.2 mg of salmon sperm DNA and 2×10^6 c.p.m. of probe per ml of hybridization solution¹⁴. The gel membranes were then washed twice at 0°C for 30 min in TMA (3 M tetramethylammonium chloride (Aldrich), 2 mM EDTA and 50 mM Tris pH 8.0), once each at 23°C (30 min) and 60°C (7 min) in TMA containing 0.2% SDS, and at 23°C for 30 min in TMA¹⁴. The gel strips were then autoradiographed between two Quanta III intensifier screens (Dupont) at -80°C.

probe (Fig. 3a, lane 5). As expected, a second unaffected sibling in this family who is heterozygous for mutant haplotype 3 and normal haplotype 4 alleles hybridized to both the mutant and normal probes (Fig. 3a; lanes 6 and 12, respectively). Thus the oligonucleotide hybridization analysis not only was able to detect the specific mutation in genomic DNA, but also demonstrated concordant segregation of the normal and mutant alleles in this PKU family.

To verify that the mutation is not a constitutive part of the normal haplotype 3 allele, we analysed a PKU family possessing

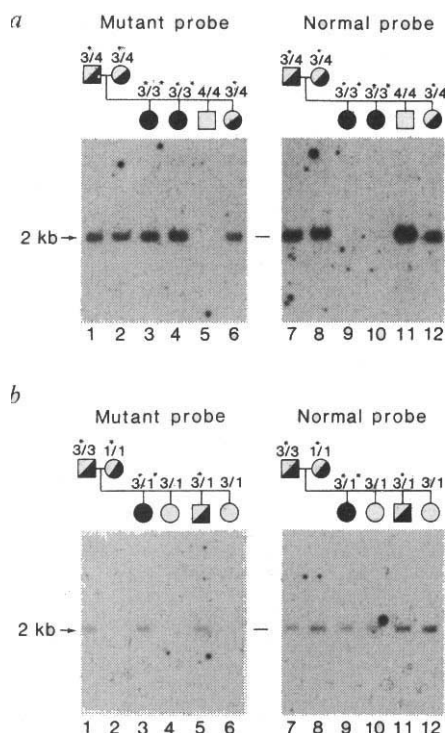


Fig. 3 Oligonucleotide hybridization analysis of two Danish PKU families. In both families, the probes used for lanes 1-6 and lanes 7-12 were the mutant and normal oligonucleotides, respectively. Squares, male; circles, female; open symbols, normal haplotype; solid symbols, mutant haplotype. **a**, Lanes 1 and 7, father; lanes 2 and 8, mother; lanes 3, 4, 9, 10, two affected children; lanes 5, 6, 11, 12, two unaffected children. **b**, lanes 1 and 7, father; lanes 2 and 8, mother; lanes 3 and 9, proband; lanes 4-6 and 10-12, three unaffected siblings. DNA was isolated from leukocytes of family members, digested with *Pvu*II and analysed by dried-gel hybridization as described in Fig. 2 legend. The segregation of the PKU alleles (*) and the appropriate RFLP haplotypes are shown at the top of each autoradiograph.

both normal and mutant haplotype 3 alleles. The father in this family, who is homozygous for haplotype 3 but an obligate carrier of the PKU trait, possesses a mutant haplotype 3 allele which hybridized to the mutant probe (Fig. 3b, lane 1) and a normal haplotype 3 allele which hybridized to the normal probe (Fig. 3b, lane 7). The results demonstrated conclusively that the splicing mutation is not a constitutive part of the normal haplotype 3 alleles *per se*. In contrast, the mother possesses normal and mutant haplotype 1 PAH alleles. The mutant probe did not hybridize to DNA isolated from this individual (Fig. 3b, lane 2), while the normal probe did (Fig. 3b, lane 8). The data indicate that the PKU mutation associated with the haplotype 1 allele in this family is not the same as that identified in the mutant haplotype 3 allele. The proband in this family, having inherited a mutant haplotype 3 allele from the father and a mutant haplotype 1 allele from the mother, hybridized to both the mutant and normal probes (Fig. 3b, lanes 3, 9). This individual apparently contains two different mutant alleles and is thus a compound heterozygote.

Carrier detection by oligonucleotide analysis

Our hybridization analysis of the above family also illustrated the feasibility of using the mutant probe to detect carriers of the PKU trait. There are three phenotypically normal siblings in addition to the proband in that family. Since both parents are haplotype homozygotes (1/1 and 3/3), all offspring are haplotype heterozygotes (1/3) and the RFLP haplotype analysis is therefore totally non-informative. The observed 'normal'

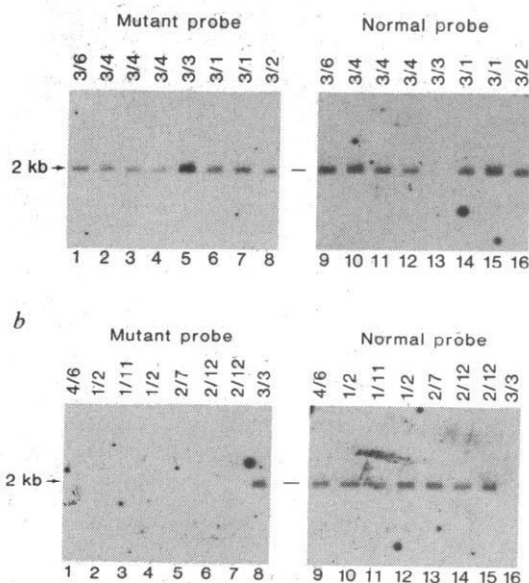


Fig. 4 Oligonucleotide hybridization analysis of Danish PKU individuals of defined RFLP haplotypes. DNA was isolated from leukocytes of PKU individuals and analysed using mutant (lanes 1-8) and normal (lanes 9-16) oligonucleotide probes as described in Fig. 2 legend. **a**, Haplotype 3 genotypes; **b**, non-haplotype 3 genotypes. The previously characterized RFLP haplotypes are shown at the top of each lane.

phenotypes in the unaffected siblings can result from three genotypes: normal haplotype 3/normal haplotype 1; mutant haplotype 3/normal haplotype 1; and mutant haplotype 1/normal haplotype 3. These genotypes can be partially distinguished by direct analysis of the splicing mutation site within the gene. DNA from one of the siblings hybridized to the mutant oligonucleotide (Fig. 3b, lane 5), suggesting that this individual has inherited the mutant haplotype 3 allele from his father and the normal haplotype 1 allele (Fig. 3b, lane 11) from his mother, and must therefore be a carrier of the PKU trait. In the absence of information for the haplotype 1 mutant gene sequence, the carrier status of the other two siblings cannot yet be fully determined. Nevertheless, the possibility that they have also inherited the mutant haplotype 3 allele from their father is excluded by the negative hybridization data obtained with the mutant oligonucleotide probe (Fig. 3b, lanes 4, 6).

Mutation frequency and haplotype association

We next determined the frequency of the splicing mutation and the degree of its association with haplotype 3 in the Danish population. RFLP haplotype analysis of 33 informative PKU families (66 normal and 66 mutant PAH chromosomes) provided a repertoire of mutant alleles for such an analysis (Table 1). Genomic DNA samples isolated from all PKU individuals containing a mutant haplotype 3 allele hybridized to the mutant probe (Fig. 4a, lanes 1-8). One PKU individual possesses two copies of the mutant haplotype 3 allele, and his DNA hybridized more strongly to the mutant probe (Fig. 4a, lane 5). The results suggested that all mutant haplotype 3 alleles in Denmark contain the same mutation in the PAH gene.

Although the homozygous haplotype 3 patient showed no hybridization with the normal probe as expected (Fig. 4a, lane 13), the normal probe did hybridize to all mutant alleles in the remaining individuals who possess non-haplotype 3 chromosomes (Fig. 4a, lanes 9-12, 14-16). Another group of PKU individuals possessing various non-haplotype 3 mutant alleles were also analysed. DNA isolated from these individuals did not hybridize to the mutant probe (Fig. 4b, lanes 1-7), while

Table 2 Association of splicing mutation with mutant haplotype 3 alleles in Denmark

Haplotypes	Mutant probe analysis*		Normal probe analysis*	
	Normal	PKU	Normal	PKU
1	0/5	0/10	5/5	10/10
2	ND	0/12	ND	12/12
3	0/2	23/23	2/2	0/23
4	0/6	0/8	6/6	8/8
5	0/1	0	1/1	0
6	0	0/2	0	2/2
7	0/2	0/1	2/2	1/1
8	0/1	0	1/1	0
9	0	0/1	0	1/1
10	0/1	0	1/1	0
11	ND	0/1	ND	1/1
12	0	0/2	0	2/2

ND, not determined due to lack of DNA samples.

* Number of hybridizing alleles/number of genes analysed.

strong hybridization signals were obtained with the normal probe (lanes 9–15). As controls in these experiments, we included DNA isolated from a homozygous haplotype 3 PKU individual; this hybridized to the mutant probe (Fig. 4b, lane 8) but not to the normal probe (lane 16). Our experiments strongly suggest that the mutations in all non-haplotype 3 alleles are distinct from the G-T to A-T transition mutation present in the mutant haplotype 3 allele.

Table 2 summarizes the population genetic data. A total of 91% (60 out of 66) of the available PKU chromosomes and 27% (18/66) of the available normal chromosomes were analysed with the mutant and normal oligonucleotide probes. All mutant haplotype 3 alleles hybridized with the mutant probe and none hybridized with the normal probe (Table 2). The remaining mutant alleles of other haplotypes and all of the normal alleles hybridized specifically with the normal probe. The results demonstrate unambiguously that the splicing mutation is specifically associated with the mutant haplotype 3 alleles, and the association is both inclusive and exclusive.

Discussion

Although PKU appears to be a heterogeneous disorder at the clinical level, controversy has persisted over whether this heterogeneity results from multiple mutations in the PAH gene varying in their degree of severity (for review see ref. 15). We recently reported that the cause of the disease must be heterogeneous because PAH mRNA was detected in some, but not all, PKU liver biopsy specimens¹⁶. Liver-specific expression of the PAH gene in man and the general lack of hepatic tissues of sufficient quantity and quality has precluded study of the molecular basis of PKU at the mRNA and protein levels. Having recently established the normal PAH gene structure⁷, however, PKU mutations can be identified at the genomic DNA level by characterization of the mutant genes.

Extensive RFLP haplotype analysis of the PAH locus in the

Danish population provided the theoretical basis for selection of mutant alleles for molecular characterization. It is interesting that 90% of the PKU genes in the Danish population are confined to four common haplotypes. We have reported here the molecular cloning of a mutant PAH gene associated with haplotype 3, which comprises 38% of all mutant alleles in the Danish population. Sequence analysis of this gene demonstrated a single base substitution at the 5' splice donor site of intron 12. Expression studies using the mutant sequence has demonstrated that the mutation causes aberrant RNA splicing by skipping exon 12 in the mature mRNA, and the mRNA is translated into a truncated protein product that is unstable in the cell (J.M. *et al.* unpublished results). DNA hybridization analysis using an oligonucleotide probe specific for the genetic lesion demonstrated an absolute association of this mutation with the mutant haplotype 3 alleles in the Danish population. The results strongly suggest that the G-T to A-T mutation occurred relatively recently on a normal haplotype 3 gene and spread in the Danish population by a founder effect. In addition, the specific mutation is not present in mutant alleles of other haplotypes, providing unambiguous evidence that there are multiple and distinct mutations in the PAH gene which are responsible for PKU.

Since PKU frequency among Caucasians is the highest in Ireland and progressively less prevalent throughout northern and southern Europe, it has been proposed that the PKU allele is of Celtic origin and spread to various European populations¹⁷. Using the mutant-specific probe to analyse DNA samples of PKU kindreds from various European countries, we have obtained preliminary evidence that the splice donor site mutation of intron 12 is also present in England, Ireland, Scotland, Switzerland and Italy (data not shown). In those families that are informative for RFLP haplotype analysis, the association between the specific mutation and RFLP haplotype 3 is also preserved in these populations. Thus, the data suggest the spread of a single mutant haplotype 3 chromosome in the Caucasian race by founder effect.

The future characterization of PKU mutations arising on the other prevalent haplotypes (haplotypes 1, 2 and 4) will clarify whether PKU is caused by a limited number of mutations that spread throughout the Caucasian race or whether multiple PKU mutations arose independently in various population backgrounds. Such a study will provide insight into the evolution and molecular origin of mutations in the PAH gene which cause PKU. If such studies confirm our hypothesis regarding the linkage of specific mutations and RFLP haplotypes in the PAH locus, it should be possible to design a cassette of oligonucleotide probes to detect 90% of the mutation chromosomes in the Caucasian population and provide a potential molecular means for PKU carrier detection in individuals with no family history of PKU.

We thank Ms Kelly Brayton for technical assistance, Mr David Nunn for the synthesis of specific oligonucleotides, and Ms Kelly Porter for typing the manuscript. This work was partially supported by NIH grant HD-17711 to S.L.C.W., who is also an Investigator of the Howard Hughes Medical Institute. A.G.D. is the recipient of NIH postdoctoral fellowship HD-06495.

Received 19 May; accepted 23 June 1986.

- Scriver, C. R. & Rosenberg, L. E. (eds) in *Amino Acid Metabolism and Its Disorders* 290–337 (Saunders, Philadelphia, 1973).
- Scriver, C. R. & Clow, C. L. *New Engl. J. Med.* **303**, 1336–1343 (1980).
- Scriver, C. R. & Clow, C. L. *A. Rev. Genet.* **14**, 179–202 (1980).
- Kwok, S. C. M., Ledley, F. D., DiLella, A. G., Robson, K. J. H. & Woo, S. L. C. *Biochemistry* **24**, 556–561 (1985).
- Woo, S. L. C., Lidsky, A. S., Güttler, F., Chandra, T. & Robson, K. J. H. *Nature* **306**, 151–155 (1983).
- Lidsky, A. S. *et al. Am. J. hum. Genet.* **37**, 619–634 (1985).
- DiLella, A. G., Kwok, S. C. M., Ledley, F. D., Marvit, J. & Woo, S. L. C. *Biochemistry* **25**, 743–749 (1986).

- Chakraborty, R. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Lidsky, A. S., Güttler, F. & Woo, S. L. C. *Lancet* **1**, 549–551 (1985).
- Orkin, S. H. & Kazanian, H. H. *A. Rev. Genet.* **18**, 131–172 (1984).
- Kidd, V. J., Wallace, R. B., Itakura, K. & Woo, S. L. C. *Nature* **304**, 230–234 (1983).
- Studnicki, A. B. & Wallace, R. B. *DNA* **3**, 7–15 (1984).
- Wood, W. I., Gitschier, J., Lasky, L. A. & Lawn, R. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1585–1588 (1985).
- DiLella, A. G. & Woo, S. L. C. *Meth. Enzym.* (in the press).
- Güttler, F. *Acta paediat., Stockh. Suppl.* **280**, 7–80 (1980).
- DiLella, A. G., Ledley, F. D., Rey, F., Munnich, A. & Woo, S. L. C. *Lancet* **1**, 160–161 (1985).
- Woolf, L. I. *J. Ir. med. Ass.* **69**, 398–401 (1976).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
- DiLella, A. G. and Woo, S. L. C. *Meth. Enzym.* (in the press).

Effect of gravitational lenses on the microwave background, and 1146+111B,C

Jeremiah P. Ostriker* & Ethan T. Vishniac†

* Princeton University Observatory, Princeton, New Jersey 08544, USA

† Astronomy Department, University of Texas, Austin, Texas 78712, USA

We suggest a simple check of the proposal by Turner *et al.*¹ that the pair of quasars 1146+111B,C are, in fact, two images of the same object lensed by an intervening object. If the lens is a cluster, either of normal galaxies or of dark matter, the microwave background seen through the cluster will be distorted by the Zeldovich-Sunyaev effect at the 10^{-3} level.

A new and exceptional candidate for an astronomical gravitational lens system has been found by Turner *et al.*¹. Two quasars separated on the sky by 157 arc s have very similar spectra and redshifts agreeing to within 100 km s^{-1} . If confirmed by further observations, this would be by far the largest angular splitting observed. Galaxies in clusters will typically produce a splitting of ~ 3 arc s (ref. 2), not far from the average of the six known systems, which is 4 arc s. As the splitting angle is proportional to $(v_{\parallel}/c)^2$, where $v_{\parallel}$ is the internal one-component velocity dispersion and c is the velocity of light, clusters of galaxies can produce far larger splittings. In fact, on this basis Paczynski³ proposed that the 1146+111B,C pair might be such a lens system before the recent observations.

This system, if it is a gravitational lens, may have been produced by a massive black hole⁴ or by a cosmic string⁵, but it is worth exploring the more conventional possibility that it arises from an intervening cluster of matter. We now show that, if a cluster has acted as a gravitational lens, it is essentially guaranteed that it will also produce an observable Zeldovich-Sunyaev effect. The new lens system corresponds to a very large and detectable effect. This is not surprising, as typical clusters can produce a nearly detectable effect, and this lens system would require an unusual cluster.

Turner *et al.*² showed that, under normal circumstances, the production of a double image requires a projected mass distribution (on the plane of the sky) greater than Σ_c , where for a universe with density parameter $\Omega_0 = 1$

$$\Sigma_c = \frac{cH_0}{8\pi G} \frac{x^2(y^{1/2}-1)}{(y^{1/2}-x^{1/2})(x^{1/2}-1)} \quad (1)$$

where $y \equiv 1+z_Q$ and $x \equiv 1+z_L$ are the redshifts of the quasar and the lens, respectively, $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the Hubble constant, and G is the universal gravitational constant. This surface density is of order 1 g cm^{-2} , and thus the optical depth to electron scattering (τ_{es}) is likely to be significant for all gravitational lens systems. If the ratio of baryonic mass to total mass is f_b , and the matter in the lens has normal composition, and is ionized, then

$$\tau_{es,c} = 0.35 f_b h \Sigma_c \quad (2)$$

so that, if $\Sigma > \Sigma_c$, then $\tau_{es} > \tau_{es,c} \approx 1$.

The angular splitting ($\Delta\theta$) for an isothermal distribution of mass is, for the chosen cosmology²,

$$\Delta\theta = (v_{\parallel}/c)^2 8\pi \frac{y^{1/2}-x^{1/2}}{x^{1/2}(y^{1/2}-1)} \quad (3)$$

For an isothermal sphere with gas in equilibrium with the same distribution as the gravitating mass, $v_{\parallel}^2 = C_s^2$, where C_s is the speed of sound, we will take as a general result that

$$C_s^2 \equiv \beta v_{\parallel}^2 \quad (4)$$

Finally, the Zeldovich-Sunyaev effect for temperature fluctuations in the microwave background can be written as

$$\Delta T/T = -2.2 \times 10^3 C_s^2 / c^2 \tau_{es} \quad (5)$$

in the long-wavelength limit.

Combining equations (1)–(5) we obtain

$$|\Delta T/T| \geq |(\Delta T/T)c| = 1.78 f_b \beta \Delta\theta h g(x, y) \quad (6)$$

where

$$g(x, y) \equiv \frac{x^{5/2}(y^{1/2}-1)^2}{(y^{1/2}-x^{1/2})^2(x^{1/2}-1)}$$

For the case at hand, $y = 2.01$. The minimal value of $g(x, y)$ occurs, for a fixed quasar position y , at $x = [y^{1/2} + 3/4 - ((4y^{1/2} + 3)^2 - 40y^{1/2})^{1/2}/4]^2$. In this case this is $x = 1.20$, which gives a value of $g(x, y) = 27.8$. Therefore, for the system 1146+111B,C where $\Delta\theta = 157$ arc s, we find

$$|\Delta T/T| \geq 0.038 \beta f_b h \quad (7)$$

From light-element nucleosynthesis, the fraction of the critical density contributed by baryons, $\Omega_b = 0.03 h^{-2}$ (ref. 6), so that if $f_b = \Omega_b$, then $|\Delta T/T| = 1.2 \times 10^{-3} \beta / h^{-2}$, which should definitely be detectable. For the more likely value of $x = 1.5$, the numerical coefficient is larger by a factor of ~ 1.77 .

This calculation has assumed that the cluster has the shape of an isothermal sphere. Is this a fair assumption? The relationship between the angular splitting and the surface density of material is relatively insensitive to the shape of the cluster. However, if the cluster is elongated towards the observer, then the integrated electron pressure is no longer simply related to the angular splitting. This is the significance of the factor β in our calculation. Assuming we are observing lensing due to a prolate spheroid pointed at us with an axis ratio < 1 , then β is roughly proportional to the value of the axis ratio. For an average lens system we can be confident that this will not greatly lower our estimate of the induced temperature distortion. For the particular case of 1146+111B,C we have no basis on which to constrain the axis ratio.

Suppose that the region has, for some reason, excluded gas, or that the gas has been kept cold, or that f_b is in fact quite small. In this case, there will be no electrons in the intracluster gas and therefore no Zeldovich-Sunyaev effect. There will still be an effect on the microwave background due to the mass concentration. This is due to purely gravitational effects on the background photons as they traverse the lens system. This effect is approximately given by $|\Delta T/T| = (\Delta\theta)^2$ (see ref. 7), which in this case gives a temperature distortion of $\sim 6 \times 10^{-7}$. This is smaller than present limits on detectability.

We see that an unusual cluster of a normal type would produce a large effect. Even if the cluster were made of dark matter without galaxies, but contained baryons in a normal ratio, the effect would be undiminished. In the unlikely event that the cluster were devoid of hot gas or of some very unusual shape so that the normal effect were absent, there would still be an effect not far below our present ability to look for temperature distortions. These results suggest that the microwave background at the lens system 1146+111B,C, and any other lens candidates with large separations, necessarily provide us with an independent way of confirming their status as gravitational lenses.

We thank Edwin L. Turner for helpful comments. This work was supported by NSF grant AST-8451736 and NASA grant NAGW-765.

Received 4 April; accepted 2 May 1986.

1. Turner, E. L. *et al. Nature* **321**, 142–144 (1986).
2. Turner, E. L., Ostriker, J. P. & Gott, J. R. *Astrophys. J.* **284**, 1–22 (1984).
3. Paczynski, B. *Nature* **319**, 567–568 (1986).
4. Paczynski, B. *Nature* **321**, 419–420 (1986).
5. Gott, J. R. *Astrophys. J.* **288**, 422–427 (1985).
6. Yang, J., Turner, M. S. & Steigman, G., Schram, D. N. & Olive, K. A. *Astrophys. J.* **281**, 493–511 (1984).
7. Rees, M. J. & Sciama, D. W. *Nature* **217**, 511–516 (1968).

Observations of the cosmic background radiation near the double quasar 1146+111B,C

Antony A. Stark, Mark Dragovan, Robert W. Wilson
& J. Richard Gott III

AT&T Bell Laboratories, Holmdel, New Jersey 07733, USA
Princeton University Observatory, Princeton, New Jersey 08544, USA

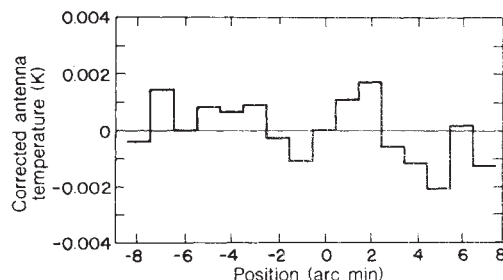


Fig. 1 Corrected antenna temperature at 100 GHz as a function of position along a constant declination strip passing between the quasars 1146+111B,C. The 0 arc min position is $\alpha_{1950} = 11^{\text{h}} 46^{\text{m}} 06^{\text{s}}$, $\delta_{1950} = 11^{\circ} 05' 28''$. West is to the left, east to the right.

Paczynski¹ and Turner *et al.*² have suggested that the extraordinary double quasar 1146+111B,C found by Hazard *et al.*³ is a possible gravitational lens. The hypothetical lensing object is unknown. If the lensing object is a cluster of galaxies⁴ or a cosmic string^{5,6}, it may cause observable effects in the cosmic background radiation. At millimetre wavelengths, a cluster would be expected to cause a reduction in the temperature of the background in the vicinity of the cluster as the result of the Zeldovich-Sunyaev effect. The area of sky on one side of a cosmic string would have a different radiation temperature from that on the other side. If either of these expected signatures were seen in the cosmic background, it would be strong evidence in favour of that hypothesis. We have searched for microwave background inhomogeneities near 1146+111B,C. An east-west strip of sky, 16 arc min long and centred on the point between the quasars, was observed with a beam of diameter 105 arc s at 3 mm wavelength. Our observations show only noise with an r.m.s. noise level of 0.0010 K. These observations set limits on the properties of the lensing object if it is a cosmic string or a cluster of galaxies, and we do not confirm either hypothesis.

Observations were made for 4-h periods on each of seven nights in March and April 1986, using the 7-m-diameter antenna at AT&T Bell Laboratories, Crawford Hill. At 100 GHz, this antenna has a beam efficiency of 0.92 and a beam size of 105 arc s (full width at half maximum). The receiver is a superconductor-insulator-superconductor (SIS) mixer followed by an FET (field-effect transistor) amplifier with 512 MHz bandwidth. Double-sideband receiver temperatures were typically 90 K, and the total system temperature including atmospheric corrections was 180–300 K. The receiver gain and noise were measured every half hour by chopping between a liquid-nitrogen-cooled absorber and a room-temperature absorber. The sky temperature was measured by chopping between the sky and the liquid nitrogen absorber, which allows a calculation of the atmospheric opacity.

The observing method was a beam-switched drift scan, as described by Radford *et al.*⁷. As the source moved across the sky, the antenna was driven to a position ahead of the source and then held stationary with respect to the Earth. The observed strip would then drift through the beam. To eliminate the effect of gain fluctuations, the receiver was chopped at a frequency of 15 Hz between the primary beam and a reference position at 30 arc min greater azimuth, so that the observed value of a point on the strip is the difference in antenna temperature between that point and the reference position. The time series that resulted from each drift scan was binned at intervals corresponding to 1 arc min on the sky, slightly larger than half a beam-width. Many drift scans were averaged to yield the plot in Fig. 1. The reference positions changed as the source moved across the sky, because of field rotation. The chopping was done with a rotating blade at the Cassegrain focus⁸. Before each night's observations the system was checked by drift-scan observations of calibration sources with known millimetre-wave fluxes.

The results are shown in Fig. 1. Here, all the drift-scans have been averaged. None of the data have been discarded. Figure 1 shows corrected antenna temperature as a function of position along the strip. The zero-point of the temperature scale has been arbitrarily set so that the average value along the strip is zero;

therefore, the data do not show possible differences between the strip and the reference positions. However, a step or a slope in the brightness as a function of position along the strip would appear in the plot.

These data seem to be random. The r.m.s. noise level is 0.0010 K. If the lensing object is an isothermal cluster there is a predicted⁴ Zeldovich-Sunyaev dip of $\Delta T \approx -0.003 K \beta f h^{-2}$ centred on the 0 arc min position in Fig. 1, where β is a geometrical factor of order unity, f is a factor of order unity that measures the fraction of the mass in the form of ionized gas ($f \equiv M_{\text{ionized}}/[0.03 M_{\text{virial}}]$; for the Coma cluster $f = 1$) and $h \equiv H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, where H_0 is the Hubble constant. If the lensing object is a straight cosmic string there is a predicted⁶ step of $\Delta T = 0.002 K v_s/c$ of either sign where v_s is the velocity of the presumably relativistic string perpendicular to its length and to our line of sight. The step could be located anywhere from position -3 to $+3$ arc min, because the string might lie anywhere between the quasars. It would be broadened by ~ 2 beam-widths because the position angle of the string is most likely to be 63° , that is, perpendicular to the line between the quasars. Both a cluster and a string would be expected to affect several of the independently measured points on our spatial strip, so the statistical significance of our non-detection is greater than the noise in any single point. The data are consistent with a limit of $\beta f h^{-2} \leq 0.3$ for the cluster. The hypothetical cluster would then have to be relatively poor in ionized gas: the factor $M_{\text{ionized}}/M_{\text{virial}}$ would be one third of the Coma cluster value. The data are also consistent with a limit of $v_s \leq 0.5 c$ (where c is the speed of light) for the case of a string. A string would be expected⁵ to have $v_s \approx 0.5 c$, so our measurement does not rule out the string hypothesis, but this measurement constrains the properties of the string. For example, the apparent projected angle, χ , between the string and the normal to the line between the quasars is given by $\tan(\chi) = (v_s/c) \tan(\alpha)$, where α is the angle between the string and the plane of the sky.

The data are, of course, also consistent with the negative hypothesis that B and C are two different quasars that coincidentally have very similar spectra, and that there is no lens at all. At the 3σ level of statistical significance, there is no effect on scales of ≥ 105 arc s at a level of 10^{-3} of the cosmic background radiation.

We thank R. E. Miller for the SIS junction used in the receiver, and W. H. Bent for contributions to instrumentation on the 7-m antenna.

Received 14 April; accepted 10 June 1986.

1. Paczynski, B. *Nature* **319**, 567–568 (1986).
2. Turner, E. L. *et al.* *Nature* **321**, 142–144 (1986).
3. Hazard, C., Arp, H. & Morton, D. C. *Nature* **282**, 271–272 (1979).
4. Ostriker, J. P. & Vishniac, E. T. *Nature* **322**, 804 (1986).
5. Gott, J. R. *Astrophys. J.* **288**, 422–427 (1985).
6. Gott, J. R. *Nature* (submitted).
7. Radford, S. J. E. *et al.* *Astrophys. J.* **300**, 159–166 (1986).
8. Goldsmith, P. F. *Bell System Tech. J.* **56**, 1483 (1977).

The chemical evolution of the Galaxy

Gerard Gilmore* & Rosemary F. G. Wyse†

* Institute of Astronomy, Madingley Road,
Cambridge CB3 0HA, UK

† Space Telescope Science Institute, Baltimore,
Maryland 21218, USA

The distribution of enriched material in the stars and gas of our Galaxy contains information pertaining to the chemical evolution of the Milky Way from its formation epoch to the present, providing general constraints on theories of galaxy formation. Detailed studies of the metallicities of well-defined samples of long-lived G-dwarf stars in the solar neighbourhood have ruled out the 'simple closed-box' model of galactic chemical evolution—too few very metal-poor stars are observed^{1–3}. The weakest assumption inherent in this model is that the galactic disk formed and evolved as a closed system. Allowing accretion of gas, either metal-enriched or primordial (the latter requiring a particular dependence of star formation on gas density), can yield an improved fit to the observations³. However, secondary infall from the galactic extreme spheroid provides only a slight alleviation of this 'G-dwarf problem'⁴. Here we show that consideration of the chemical properties of the thick-disk population of the Galaxy^{5,6} results in a self-consistent model for galactic chemical evolution.

Motivation for the closed-box model for the galactic disk came from the proposal⁷ that the formation of the spheroid of the Galaxy, and the collapse of residual gas to a thin disk, was extremely rapid, having been largely completed in a free-fall time. However, recent evidence from *in situ* observations of a stellar component of the Galaxy with kinematics and metallicity distribution intermediate between those of the extreme spheroid and thin disk^{5,6,8–11} has shown that this picture of Galaxy formation and evolution is oversimplified. Here we investigate the metal-enrichment history of an evolutionary sequence in which a short-lived phase of rapid star formation forms the extended, metal-poor, pressure-supported extreme spheroid⁷, with the thick disk forming subsequently before a final equilibrium state is attained, the centrifugally supported thin disk forming over a long timescale once the Galaxy potential is unperturbed and steady^{6,12,13}.

The transition between extreme spheroid and thick disk can be investigated using the two-zone model⁴. Thus, we assume that gas is removed from the spheroid star formation process at a rate which is a constant multiple, denoted by c , of the star formation rate. This may be expected if the gas is heated by supernova explosions, and within the context of the modified simple model has the same effect as a lowered nucleosynthetic yield⁴. We keep the remaining major assumptions of the simple model; namely, that the stars enrich the gas instantaneously and that the system is well mixed and chemically homogeneous. The parameters in the model have standard values: solar neighbourhood yield¹⁴ and stellar initial mass function¹⁵, with a lock-up fraction of gas into stars of 0.6–0.8 (ref. 16). For the extreme spheroid, we also keep the assumption of zero initial metallicity, in keeping with the most recent derivation of the metallicity distribution of extremely metal-poor spheroid stars¹⁷.

The theoretical cumulative metallicity distribution, whose form in the adopted model is independent of the unknown star formation rate, may be compared⁴ with observations to solve for the value of the constant c . We have used the metal-poor galactic globular clusters⁸ to define the observed extreme-spheroid metallicity distribution (Fig. 1a). This distribution is in excellent agreement with that derived by Sandage¹⁰ for his

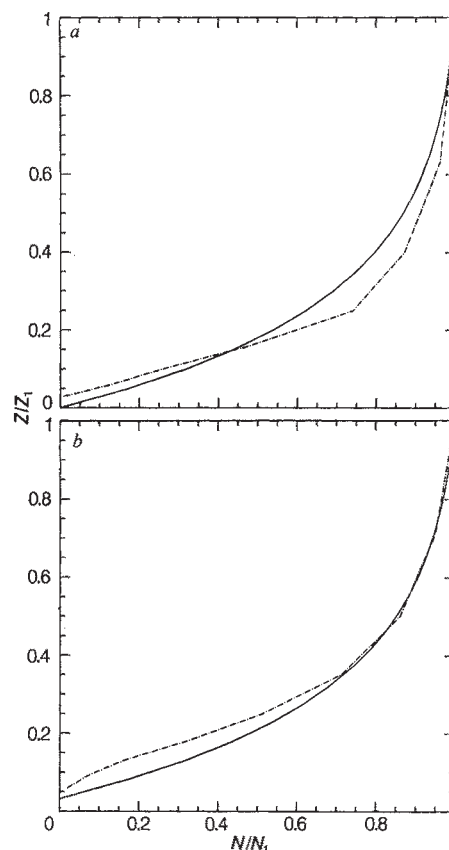


Fig. 1 Comparison of observed (dashed lines) and predicted (solid lines) cumulative metallicity distributions for the extreme spheroid⁸ (a) and thick disk^{6,10,11} (b). Plots show the total number, N , of tracers with metallicity $\leq Z$, normalized to the maximum abundance, Z_i . In a, the observed distribution⁸ is defined by the metal-poor ($[\text{Fe}/\text{H}] < -0.8$) galactic globular clusters; the model shown has $c = 10$. In b, the observed log metallicity distribution has been approximated by a truncated gaussian ($\pm 2\sigma$), with mean $[\text{Fe}/\text{H}] = -0.6$, $\sigma = 0.3$; the model shown has $c = 1$.

large, kinematically defined sample of subdwarf stars. Norris¹¹ has shown that this distribution represents the true metallicity structure of the extreme spheroid, independent of the detailed selection criterion for the sample. For the adopted stellar nucleosynthesis parameters, these data are fit well by removal of gas from the star-forming process at a rate ~ 5 –20 times the spheroid star formation rate (see refs 4, 18, 19). The uncertainty in c results from the possible range of the lock-up fraction in the extreme spheroid, and from the observational spread; the use of the cumulative metallicity distribution forces a smooth distribution through the observations. These values of c yield an average metallicity of the residual gas of $[\text{Fe}/\text{H}] = -1.4 \pm 0.2$. (As the gas removed from the spheroid star-forming process is, as usual, assumed to be well mixed and enriched with stellar ejecta with a time-independent, solar distribution of elements, the time-averaged metallicity is most meaningful.) This metallicity is in excellent agreement with the metal-poor tail of the thick-disk stars^{6,9,11}, but is much too metal-poor for the thin-disk stars—showing explicitly that an initial enrichment due to rapid infall from the extreme spheroid cannot alone provide a resolution of the thin-disk G-dwarf problem⁴.

In the present model, the residual gas after spheroid star formation ceases is treated as simply providing the proto-thick disk with a non-zero initial metallicity, equal to the mean gas enrichment found above. This treatment is consistent with the high star formation rate suggested by the extreme-spheroid kinematics and the rapid removal of all gas from the spheroid

† Present address: University of California, Berkeley, California 94720, USA.

star-formation process in the present model ($c = 5-20$). As the thick-disk metallicity distribution overlaps that of the extreme spheroid at the metal-poor end, and that of the thin disk at the metal-rich end, one must be careful to ensure minimal inter-component contaminations. The thick disk may be defined in three ways, the first by a volume sample *in situ* above the galactic plane, where it is expected to dominate⁶, the second kinematically¹⁰, and the third by a combination of spectroscopic and kinematic data¹¹. These three determinations are in good agreement, so that the adopted metallicity structure is not sensitive to a specific criterion. Additionally, the most recent analogous determinations of the thin-disk metallicity structure^{6,10,20} show that there is no difficulty in isolating the thick- and thin-disk parameters with sufficient precision for the present problem.

Figure 1b shows the cumulative metallicity distribution for thick-disk stars^{6,10,11}. The best-fitting theoretical curves have net outflow parameter $c = 1-2$, and a corresponding mean value of the metallicity of the ejecta from thick to thin disks of $[Fe/H] = -0.6 \pm 0.1$. This level of pre-enrichment agrees well with the metallicity determined for the oldest, most metal-poor G-dwarfs^{3,6}. Hence, inclusion of the thick disk in this simple model resolves the G-dwarf problem for thin-disk stars in the solar neighbourhood.

Pre-enrichment of the thin disk was recognized previously^{3,16} as a resolution by itself to the G-dwarf problem. For the present model, detailed calculations show that the cumulative metallicity distribution for thin-disk stars limits any infall of un-enriched material following the prompt enrichment by the thick-disk ejecta to a rate of <0.2 times the star formation rate, assuming constant relative rates, as may be expected¹⁴. This is in agreement with limits on present-day high-velocity infall²¹. Constraints on our model from other quantities, such as the age-metallicity relation and the age-velocity dispersion relation for long-lived thin-disk stars, are weak, as the two recent calibrations of these relations^{12,20} are in significant disagreement. The apparently extremely-high-metallicity K-giants found in the direction of the galactic centre²² are difficult to understand if they are members of the extreme-spheroid population, unless the gas which formed the central regions of the spheroid was decoupled from the rest, but their spatial frequency and velocity dispersion are, in fact, compatible with their being disk stars.

The ratio of the mass of the second zone (thick plus thin disks), to the mass of the first zone (extreme spheroid) is approximately c/α , where α is the lock-up fraction in the first zone. Adopting a value of 0.8 for α , we then predict a present-day mass ratio, between the extreme spheroid and thick and thin disks together, of $\sim 1:15$. The thick- to thin-disk transition as modelled here predicts a further decomposition of masses to $\sim 1:3:12$ for the ratios extreme spheroid:thick disk:thin disk. The later infall of primordial material, expected in most theories of galaxy formation and suggested by the thin disk metallicity distribution, will modify these ratios somewhat.

The global parameters of the Galaxy, which are required to estimate luminosities to compare with these predictions, are as

yet ill-determined. In particular, neither the radial scale-lengths nor axial ratios of any of the three major components of the Galaxy have been measured unambiguously. The most reliable technique available assumes that our Galaxy is extremely similar to other well-studied Sbc disk galaxies²³. This gives a decomposition of our Galaxy into extreme spheroid, thick disk and old thin disk²³, which implies a value of 1:11 for the ratio of the J-band luminosity of the extreme spheroid to that of the sum of the disks. Assuming similar mass-to-light ratios (~ 6), as for SO to Sbc galaxies in the B-band²⁴, yields good agreement with the value predicted above. The further decomposition into extreme spheroid:thick disk:old thin disk yields luminosity ratios 1:0.2:11, suggesting a less massive thick-disk component than predicted here. (Note that the young thin disk has negligible mass compared to the old thin disk.) However, these observational estimates for the thick disk are highly uncertain, being based in part on an extrapolation²³ of star counts in a single field to determine the global parameters. Indeed, Sandage¹⁰ finds the galactic thick disk to have a local normalization higher by a factor of $\sim 3-5$ than the earlier determination²³; this new value yields adequate agreement with the present model. By comparison, in NGC891, thought to be extremely similar to our Galaxy, the thick disk and extreme spheroid are of comparable mass²⁵. Even the $r^{1/4}$ extreme-spheroid component in our Galaxy is poorly constrained by available data, two recent analyses differing by a factor of 3.45 in the luminosity of this component^{23,25}. Improved determinations of both the metallicity and spatial-structural parameters will only be available once the major statistical surveys underway by ourselves and others are completed.

The chronological sequence of formation of the three components proposed here may be tested by study of their detailed elemental and isotopic abundances²⁶. The transitions between components are likely to occur on timescales long compared with the lifetimes of the massive stars which explode as type II supernovae, but of the order of the lifetimes of the stars that produce carbon, nitrogen and type I supernovae. The available data on the variation of oxygen (from massive stars) and iron (from intermediate-mass stars) in stars over a range of metallicities spanning extreme-spheroid to thin-disk^{27,28} show a feature near $[Fe/H] \approx -1$, the transition metallicity between the extreme spheroid and thick disk. More exact location of this feature will allow us to estimate the duration of the formation of the extreme spheroid.

Although a full analysis of the chemical evolution of the Galaxy involves too many parameters for an analytical understanding, we believe that the simple model presented here offers new insight into the formation and evolution of the Galaxy. The galactic thick disk has a significant influence on the chemical evolution of the other components of the Milky Way.

We thank NATO for a travel grant (490/84) to aid collaborative research, the American Astronomical Society for grants (to R.F.G.W.) from the Small Research Grants Program, and Jim Truran for helpful comments.

Received 10 April, accepted 20 June 1986.

- van den Bergh, S. *Astr. J.* **67**, 486-490 (1962).
- Schmidt, M. *Astrophys. J.* **137**, 758-769 (1963).
- Pagel, B. E. J. & Patchett, B. E. *Mon. Not. R. astr. Soc.* **172**, 13-40 (1975).
- Hartwick, F. D. A. *Astrophys. J.* **209**, 418-423 (1976).
- Gilmore, G. & Reid, I. N. *Mon. Not. R. astr. Soc.* **202**, 1025-1047 (1983).
- Gilmore, G. & Wyse, R. F. G. *Astr. J.* **90**, 2015-2026 (1985).
- Eggen, O. J., Lynden-Bell, D. & Sandage, A. R. *Astrophys. J.* **136**, 748-766 (1962).
- Zinn, R. *Astrophys. J.* **293**, 424-444 (1985).
- Norris, J., Bessel, M. S. & Pickles, A. J. *Astrophys. J. Suppl. Ser.* **58**, 463-492 (1985).
- Sandage, A. R. in *Stellar Populations* (eds Renzini, A., Tosi, M. & Norman, C.) (Cambridge University Press, in the press).
- Norris, J. *Astrophys. J. Suppl. Ser.* (in the press).
- Jones, B. J. T. & Wyse, R. F. G. *Astr. Astrophys.* **120**, 165-180 (1983).
- Gilmore, G. *Mon. Not. R. astr. Soc.* **207**, 223-240 (1984).
- Twarog, B. A. *Astrophys. J.* **242**, 242-259 (1980).

- Scalo, J. *Fundam. Cosmic. Phys.* (in the press).
- Tinsley, B. M. *Fundam. Cosmic. Phys.* **5**, 287-342 (1980).
- Beers, T. C., Preston, G. W. & Shectman, S. A. *Astrophys. J.* (in the press).
- Hartwick, F. D. A. *Memorie Soc. astr. ital.* **54**, 51-64 (1983).
- Fall, S. M. in *The Milky Way Galaxy* (eds van Woerden, H., Allen, R. J. & Butler, W.) 603-610 (Reidel, Dordrecht, 1985).
- Carlberg, R. G., Dawson, P. C., Hsu, T. & VandenBerg, D. A. *Astrophys. J.* **294**, 674-681 (1985).
- Mirabel, I. F. & Morras, R. *Astrophys. J.* **279**, 86-92 (1984).
- Whitford, A. E. & Rich, R. M. *Astrophys. J.* **274**, 723-732 (1983).
- Van der Kruit, P. C. *Astr. Astrophys.* **140**, 470-475 (1984).
- Faber, S. M. & Gallagher, J. S. A. *Rev. Astr. Astrophys.* **17**, 135-187 (1979).
- de Vaucouleurs, G. & Pence, W. *Astr. J.* **83**, 1163-1173 (1978).
- Truran, J. in *Stellar Populations* (eds Renzini, A., Tosi, M. & Norman, C.) (Cambridge University Press, in the press).
- Snedden, C., Lambert, D. L. & Whitaker, R. W. *Astrophys. J.* **234**, 964-972 (1979).
- Clegg, R. E. S., Lambert, D. L. & Tomkin, J. *Astrophys. J.* **250**, 262-275 (1981).

Nimbus 7 satellite measurements of the springtime Antarctic ozone decrease

R. S. Stolarski, A. J. Krueger, M. R. Schoeberl,
R. D. McPeters, P. A. Newman & J. C. Alpert*

NASA/Goddard Space Flight Center, Laboratory for Atmospheres,
Greenbelt, Maryland 20771, USA

Farman *et al.*¹ have reported a rapid decrease, since their measurements started in 1957, of the total column amount of ozone in late winter and early spring over the Halley Bay station in Antarctica (76° S, 27° W). The decrease was most pronounced in October, early spring in the Southern Hemisphere. They attributed the decrease to the increase in stratospheric chlorine due to chlorofluorocarbon release, and proposed that the unique conditions of extreme cold and low sunlight in the Antarctic winter and spring enhanced the effect. We report measurements from the Solar Backscatter Ultraviolet (SBUV) instrument and the Total Ozone Mapping Spectrometer (TOMS) aboard the Nimbus 7 satellite, a Sun-synchronous polar-orbiting satellite which passes any given point on the dayside near local noon. These provide global measurements of ozone from November 1978 to the present which confirm the reported decline of total ozone and show the phenomenon to be regional in extent. The decrease occurs during September as the Sun rises, reaching a minimum in mid-October. Seven years (1979–1985) of October monthly means show a 40% decrease in the ozone minimum and a 20% decrease in the surrounding ozone maximum.

The SBUV is a nadir-viewing instrument which obtains total ozone from measurements of the backscattered solar ultraviolet radiation in the wavelength range 310–340 nm (refs 2, 3). Measurements in the 250–310 nm range by the SBUV instrument provide information on the distribution of ozone with altitude. The TOMS instrument is designed to measure the spatial distribution of total ozone by scanning across the track of the satellite to obtain data between successive satellite orbital tracks. Both instruments make total ozone determinations using a differential absorption method which tends to cancel effects of neutral absorbers or scatters. Thus, stratospheric aerosol, such as polar stratospheric clouds (PSC), produce only second-order errors in retrievals. Errors due to wavelength-dependent scattering exist at large solar zenith angles⁴; however, PSCs are most common in winter, and vanish as the stratosphere warms in the spring^{5,6}, the time of the maximum decrease in ozone. Instrumental interference from PSCs, therefore, cannot account for the large observed ozone decrease.

One of the major advantages of the TOMS instrument is its ability to map total ozone over the entire globe on a daily basis. Figure 1 shows a daily sequence of total ozone observations for October 1984. The low total ozone over Halley Bay (indicated by an asterisk in Fig. 1) is part of a larger, elliptically shaped minimum region extending out to ~60–70° S. This region, which rotates about the pole with an irregular period of ~7–10 days, is bounded by a steep gradient where total ozone increases to values exceeding 300 Dobson units (1 DU = 10⁻³ atm cm). At 50–60° S total ozone reaches a maximum, which exhibits larger variability due to travelling planetary waves. The polar minimum and mid-latitude maximum in total ozone are features of the normal, long-term climatology of the Southern Hemisphere⁷.

The 12 panels in Fig. 1 show 12 consecutive days, from 11 to 22 October 1984. This is the time period for which the lowest values are generally recorded. Each panel is centred about the South Pole and includes data out to 45° latitude. On 11 October, Fig. 1 shows that the minimum in total ozone is relatively symmetrical about the pole and the maximum in total ozone is

located in the lower right, between ~90 and 180° E. On succeeding days the total ozone maximum region rotates clockwise, while the minimum region elongates and begins to co-rotate with the maximum. On 16 and 17 October the maximum region has reached the lower left-hand corner, between 180 and 270° E, where it seems to dissipate and subsequently reappear in the upper right-hand corner. From there it continues to move around the pole but does not pass over Halley Bay. By 20 October the ozone distribution has returned approximately to the original situation on 11 October, giving a rotation period at this time of 9 days. The next 2 days are near repeats of 12 and 13 October.

Total ozone is proportional to the pressure-weighted integral of the ozone mixing ratio, and fluctuations in the amount of total ozone may indicate a relative increase in altitude of the mixing ratio distribution, a chemical loss in the lower stratosphere, advection of ozone into a region from other locations, or all of the above. Comparisons of National Meteorological Center (NMC) 50-mbar temperatures and TOMS total ozone reveals a high degree of spatial correlation on all of the days that have been examined. This correlation can be understood as resulting from the adiabatic ascent and descent of air as it moves around the vortex. Air which descends tends to increase total ozone while warming adiabatically. The reverse occurs when air ascends.

The isobars at middle latitudes in the lower stratosphere are generally not parallel to the isotherms. As the isobars approximate the streamlines for air parcel motion, air parcels will experience large temperature excursions (20–30 K) as they circulate through the polar vortex. These excursions could play an important role in the chemistry of the air parcels and should not be neglected in model simulations. We have also noted that the streamlines pass through the edges of both the high and low total ozone regions. This suggests that the observed decrease in both the total ozone maxima and minima result from the year-to-year decrease in the mixing ratio of the group of parcels moving between the two regions. Thus, the total ozone maxima and minima decreases do not necessarily involve separate processes.

Evidence that the long-term total ozone decrease observed at Halley Bay is not unique to Halley Bay is shown in Table 1,

Table 1 Zonal mean, local minimum and local maximum total ozone values (DU) for October at 70–80° S latitude

Year	Zonal mean	Local minimum	Local maximum
1970	306	240	484
1971	334	249	482
1972	337	237	539
1979	333	235	515
1980	270	212	467
1981	266	206	422
1982	283	186	494
1983	245	166	479
1984	240	162	446

Data from Nimbus 4 BUV (1970–72) and Nimbus 7 SBUV (1979–84).

which lists the zonal mean, local minimum and local maximum total ozone values for 70–80° S latitude, for October 1970–72 from BUV and October 1979–84 from SBUV. Data from the Nimbus 4 BUV instrument, the predecessor to SBUV, were fairly complete for the 1970–72 period; in later years, coverage of the Antarctic by BUV was reduced. The BUV ozone data have been re-processed with the algorithm used for SBUV and with compatible instrument characterization functions, making it reasonable to compare the BUV total ozone with the SBUV total ozone. Antarctic total ozone in 1979 appears comparable to that observed in the early 1970s, indicating that most of the total ozone decrease has occurred since 1979, in agreement with the Halley Bay observations.

* Present address NOAA/NMC, Camp Springs, Maryland 20723, USA.

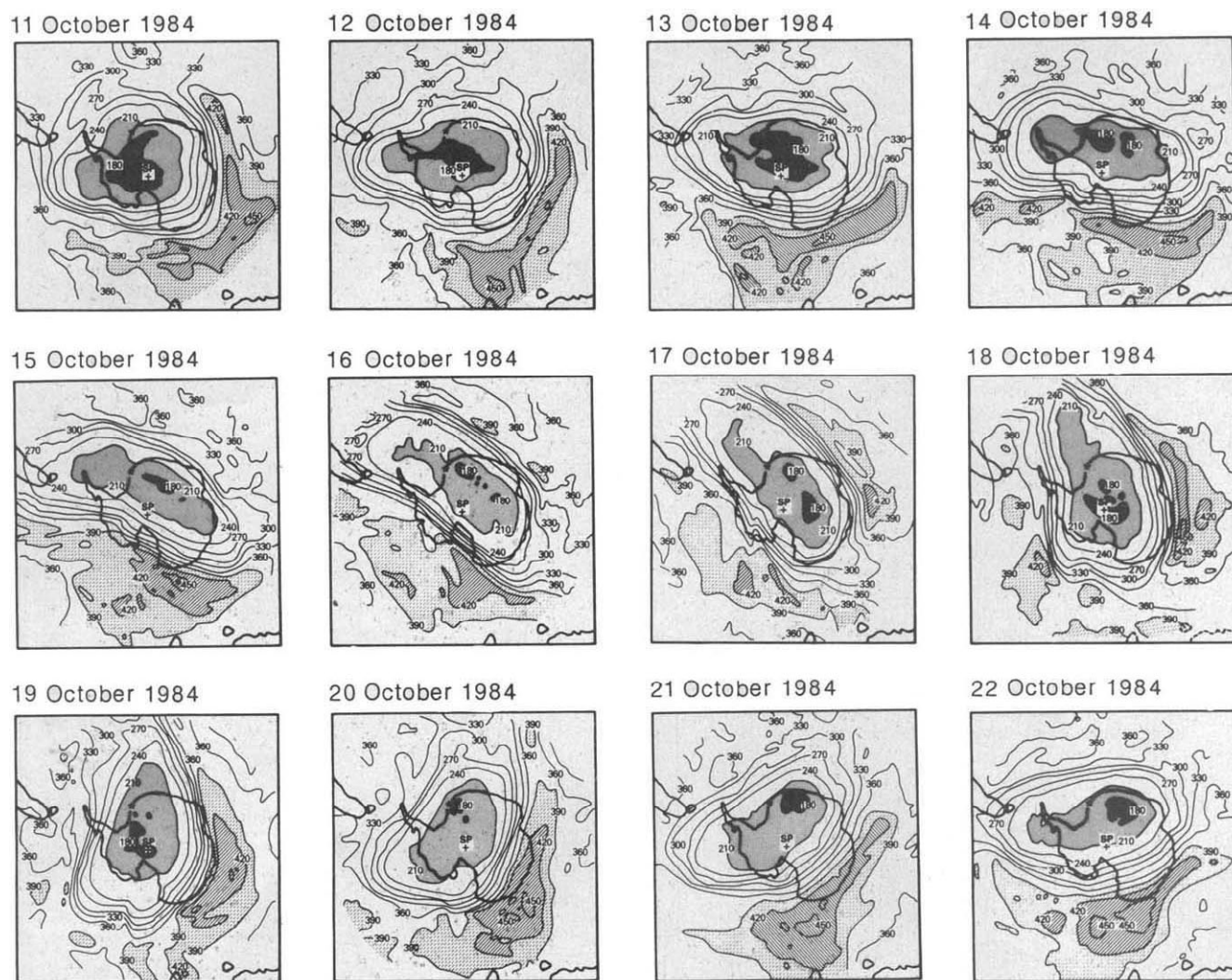


Fig. 1 Twelve-day sequence (11–22 October 1984) of TOMS measurements of total ozone content. The data are shown in south-polar projections, with the pole indicated by a cross (SP) and Halley Bay shown by an asterisk. Contours are every 30 Dobson units (1 DU = 10^{-3} atm cm). The region shown extends to $\sim 45^\circ$ latitude and the Greenwich meridian is towards the top of each diagram. Shaded regions indicate total ozone values <180 and 210 DU and >390 and 420 DU.

The nature of the decrease is shown more clearly in Fig. 2, which shows October monthly mean TOMS total ozone maps for the Southern Hemisphere in each of the seven years from 1979 to 1985. Table 2 lists the October monthly means over

Table 2 October mean ozone column over three Antarctic stations, along with the minimum and maximum values of the monthly mean (all in DU)

Year	Halley Bay (76° S, 27° W)	Syowa (69° S, 39° E)	Amundsen-Scott (90° S)	TOMS minimum	TOMS maximum
1979	273 ± 19	367 ± 47	266 ± 25	259	458
1980	232 ± 11	276 ± 22	218 ± 9	215	433
1981	244 ± 12	312 ± 30	223 ± 9	219	449
1982	221 ± 12	240 ± 49	220 ± 22	205	432
1983	199 ± 13	239 ± 32	182 ± 8	178	414
1984	190 ± 8	245 ± 24	187 ± 14	181	404
1985				152	351

Data from Nimbus 7 TOMS. Stated ranges are 1σ variances of the daily values from the monthly mean for each location.

three stations (Halley Bay, Syowa and Amundsen-Scott), as well as the minimum and maximum values of the monthly mean. The monthly mean values in the minimum region decrease from 259 DU in 1979 to 152 DU in 1985. The total ozone values for Halley Bay agree closely with those reported by Farman *et al.*¹

We have also examined four years (1979–82) of NMC 50-mbar temperature data. The interannual variations in temperature were similar to the total ozone variations, but no conclusive trend could be determined from such a short data set.

The phase of the maximum in the October monthly mean total ozone at $\sim 60^\circ$ S latitude (Fig. 2) shifts from year to year, ranging from ~ 90 to 180° E. In 1979 the maximum monthly mean total ozone value exceeds 450 DU. By 1985 the maximum is barely >350 DU, a decrease of $>20\%$ in seven years. There also seems to be a biennial modulation of the decrease in both the maximum and minimum. A quasi-biennial oscillation in total ozone has been observed previously, but with a magnitude of only ~ 6 – 8 DU (refs 8, 9).

The data given by Farman *et al.*¹ for Halley Bay showed total ozone amounts at the end of the polar night in 1984 which were a few tens of per cent lower than those at the start of the polar night. The ozone algorithms for both TOMS and SBV have been carefully scrutinized for their behaviour at high solar zenith angles and return accurate results for zenith angles up to 85° (ref. 3 and P. K. Bhartia, personal communication). We thus have observations back through the month of September for the latitude of Halley Bay (76° S). The results show values of total ozone as this region leaves polar night which are very close to those observed entering the polar night. This result is also in

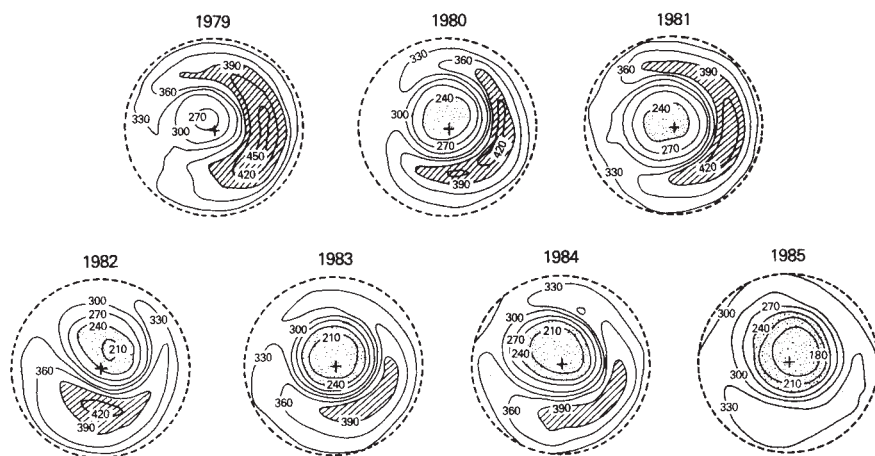
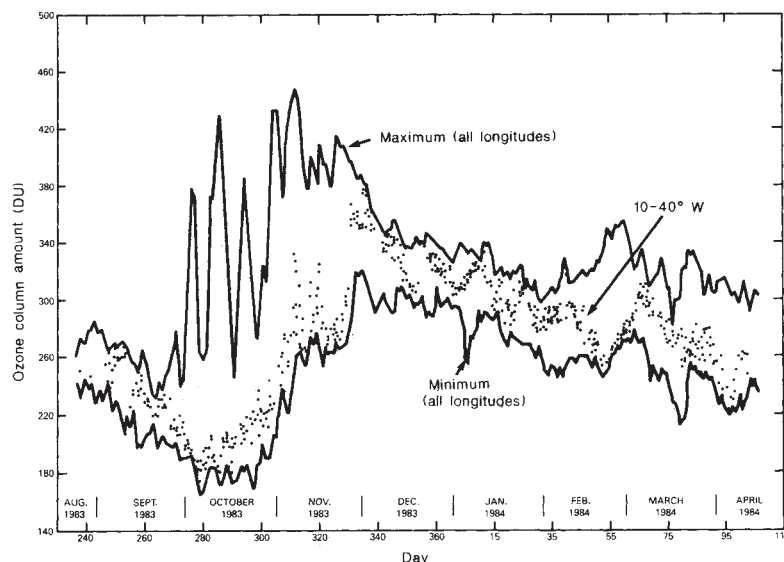


Fig. 2 Six-year sequence of October monthly means of total ozone. South polar projections, with the pole indicated by a cross and 30°S latitude by a dashed circle. The Greenwich meridian is towards the top of each panel. Contours are every 30 DU. The shaded regions indicate monthly mean total ozone amounts of <240 DU and >390 DU.

Fig. 3 Nimbus 7 SBUV data for total ozone content between latitudes 74°S and 78°S for September 1983 to April 1984. Dots, data within 15° longitude of Halley Bay (10–40°W). Solid lines, maximum and minimum values found around the entire latitude circle.



agreement with the observations of total ozone made at Syowa station (69°S) in 1982¹⁰. However, there is some indication from lunar Dobson and ozone-sonde measurements at Syowa that total ozone increases slowly until winter solstice, then decreases again until sunrise, with a more rapid decrease thereafter¹⁰.

Throughout the month of September and the first week in October, a rapid decline is observed in total ozone by both SBUV and TOMS, until the previously mentioned low values are obtained. Figure 3 shows this pattern in data from the SBUV instrument for late August 1983 to April 1984. The data points are the satellite data that are centred within $\pm 2^\circ$ latitude and $\pm 15^\circ$ longitude of Halley Bay. The solid lines are the maximum and minimum of values reported at all longitudes within the latitude band 74–78°S. Throughout October and November, the data over Halley Bay are consistently near the minimum values for all longitudes for this and other years. The decline in the minimum value of ozone column extends from late August to early October at a rate of $\sim 0.6\%$ per day. At longitudes on the other side of the pole from Halley Bay (near the total ozone maximum), the total ozone amount is more variable.

Near the end of November the low values have disappeared completely and the minimum curve attains its largest values. The jump in total ozone occurs several weeks after the final warming in the Southern Hemisphere, and marks the switch from winter to summer circulation. A similar process occurs in the Northern Hemisphere, but usually only three months after the winter solstice rather than five months, as seen here^{11,12}. Throughout the Antarctic summer the ozone column decreases again, and the Halley Bay values oscillate between the maximum

and minimum curves as lower stratospheric travelling waves pass over the station.

Our data place significant constraints on possible mechanisms for explaining the decrease. The deep minimum, or hole, follows the polar vortex, and its position is well correlated with the temperature minimum in the lower stratosphere. There also appears to be a smaller but significant decrease ($\sim 20\%$) in the total ozone maximum surrounding the polar region over the seven years studied. The decrease in the total ozone maximum is related to the decrease in the total ozone minimum, in that the streamlines indicate that air in the maximum region has passed through the periphery of the low-temperature region which is associated with the total ozone minimum.

An important constraint placed on theoretical models by these data is that the decrease in total ozone near the pole appears to take place largely in September, during twilight, not in the polar night. The maximum rate of decrease in total ozone occurs after most of the polar region is sunlit, and in 1983 was $\sim 0.6\%$ per day, extended over 40–50 days.

Any explanation of the total ozone change must be consistent with the year-to-year changes in the polar vortex, as well as with the rate of decrease through September. Various ideas have been proposed, involving combinations of chlorine chemistry, heterogeneous chemistry taking place in polar stratospheric clouds, bromine chemistry and dynamical elevation of the polar stratosphere^{1,13–15}.

We estimate that in order for chlorine chemistry at 1983 concentrations to cause the 0.6% per day decline in September, virtually all of the chlorine must be in its active state, that is,

Cl and ClO, with little or no interference from NO_x . Therefore, any proposed chlorine mechanism must be able to remove most of the chlorine from both the HCl and ClONO_2 reservoirs and to tie up NO_x , probably as HNO_3 . Such a mechanism could involve the cold temperatures and/or polar stratospheric clouds that form within the polar vortex⁶.

Any conclusions concerning the implications of the observed Antarctic decreases in total ozone for predictions of the effects of chlorine from chlorofluorocarbons must await a proven mechanism and continued observations to verify the persistence of the phenomenon. Only then will we be able to evaluate clearly the relative roles of chemistry, radiation and dynamics in contributing to the observed decrease.

We thank the Nimbus 7 ozone processing team (A. Fleig, chairman) for making available TOMS data from October 1983, 1984 and 1985; the Real-Time TOMS Project for data for a few days in October of 1985 before regularly scheduled processing; Reggie Galimore for image processing which helped make the data readily available for visual examination; S. Solomon, S. Wofsy and N. D. Sze for discussing and making available preprints of their papers; S. Chandra for useful discussions; and Lori Winter and Roberta Duffy for typing the many versions of this manuscript.

Received 8 April; accepted 1 July 1986.

- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Heath, D. F., Krueger, A. J., Roeder, H. A. & Henderson, B. D. *Opt. Engng* **14**, 323–330 (1975).
- Klenk, K. F. *et al. J. appl. Met.* **21**, 1672–1684 (1982).
- Dave, J. V. *J. Atmos. Sci.* **35**, 889–911 (1978).
- Iwasaka, Y. *et al. Middle Atmosphere Program Handbook*, Vol. 18 (ed. Kato, S.) 450–452 (SCOSTEP Secretariat, Urbana, 1985).
- McCormick, M. P., Steele, H. M., Hamill, P., Chu, W. P. & Swisler, T. J. *J. Atmos. Sci.* **39**, 1387–1397 (1982).
- Sticksel, P. R. *Mon. Weath. Rev.* **98**, 787–788 (1970).
- Hasebe, F. in *Dynamics of the Middle Atmosphere* (eds Holton, J. R. & Matsuno, T.) 445–464 (Terra Scientific, Tokyo, 1984).
- Hilsenrath, E. & Schlesinger, B. M. *J. geophys. Res.* **86**, 12087–12096 (1981).
- Chubachi, S. in *Atmospheric Ozone* (eds Zerefos, C. S. & Ghazi, A.) (Reidel, Dordrecht, 1984).
- Miller, A. J., Finger, F. G. & Gelman, M. E. *NASA tech. Memo. No. TMX-2109* (Washington, DC, 1970).
- Rood, R. B. *Pure appl. Geophys.* **121**, 1049–1064 (1983).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1986).
- Tung, K. K., Ko, M. K. W., Rodriguez, J. M. & Sze, N. D. *Nature* **322**, 811 (1986).

Are Antarctic ozone variations a manifestation of dynamics or chemistry?

Ka-Kit Tung*, Malcolm K. W. Ko†, José M. Rodriguez† & Nien Dak Sze†

* Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Clarkson University, Potsdam, New York 13767, USA

† Atmospheric and Environmental Research, Inc., Cambridge, Massachusetts 02139, USA

Observations^{1–3} reveal a large seasonal decrease in the column density of ozone during Antarctic early spring, followed by a rapid increase after October beyond its pre-spring value. Given the unique circumstances that exist in the Antarctic environment—relatively stable circumpolar vortex, cold air temperature achieved during polar night, and the increase in absorption of solar radiation by ozone as the Sun returns—we surmise the existence of a reverse circulation cell with rising motion in the polar lower stratosphere. The upwelling brings ozone-poor air from below 100 mbar to the stratosphere, possibly contributing to the observed ozone decline in early spring. At the same time, the Antarctic stratosphere might contain a very low concentration (<0.1 p.p.b.v. (parts per 10^9 by volume) of $\text{NO}_x(\text{NO} + \text{NO}_2)$, a condition that could favour a greatly enhanced catalytic removal of O_3 by halogen species^{2,4,5}.

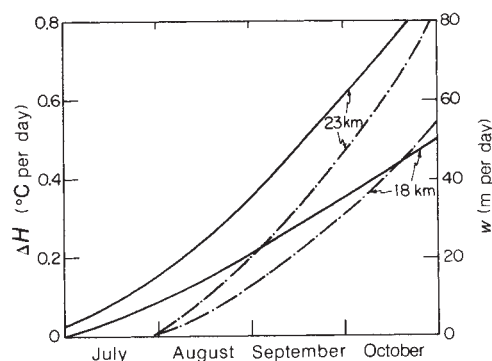


Fig. 1 Calculated heating rate (ΔH) from O_3 absorption over a 24-h period, as a function of time, for 69°S (—) and 76°S (---) at different altitudes. The heating rates are calculated from the local O_3 concentration using the parameterization of Strobel³¹. The time variation mainly reflects the seasonal changes in the number of sunlit hours. The observed O_3 profile from Syowa Station in July is adopted as the initial condition. For this calculation, horizontal transport and photochemical removal are ignored. The O_3 profile is updated in time by accounting for the effect of the upward advection $\bar{w}$ by equation (4). The right-hand scale indicates the approximate value of $\bar{w}$, assuming that $\Gamma = 10^\circ\text{C km}^{-1}$.

We argue that heterogeneous processes and formation of OCIO by the reaction $\text{BrO} + \text{ClO} \rightarrow \text{OCIO} + \text{Br}$ before and after the polar night might help to suppress the NO_x levels during the early spring period. However, the dilution of the concentrations of the chlorine species by the upwelling may reduce the effectiveness of the photochemical removal of O_3 .

The zonal mean transport in the lower stratosphere normally takes the form of a two-cell diabatic circulation, with rising motion in the tropics and subsidence near both poles^{6–10}. For a species such as O_3 , whose source is in the stratosphere, upward transport tends to reduce its column abundance, as tropospheric air with low O_3 mixing ratio is brought up to the stratosphere. Poleward and downward transport in the subsiding branches tends to increase its abundance at high latitudes. This situation is further accentuated in winter by the downward branch of the one-cell circulation that exists above 25 km (refs 11, 12). The latitudinal distribution of the column abundance of O_3 does not, however, have north-south symmetry^{13,14}, as the spring maximum in the Southern Hemisphere is located closer to the Equator than in the north.

The winter polar vortex that encircles Antarctica is less disturbed than its counterpart in the Northern Hemisphere because of the weaker stationary planetary wave perturbations forced by large-scale continental elevations and land-sea contrasts. Although the vortex may shift in position and deform in shape in response to wave perturbations, its integrity (as defined in ref. 15) is usually maintained until its breakdown in early November. It therefore appears that irreversible wave-induced transports¹⁶ do not effectively reach Antarctica during this period.

During the polar winter in the Southern Hemisphere, dynamical transport of heat into the circumpolar vortex is weak. Local temperatures in the Antarctic lower stratosphere have time to adjust so that cooling, $C(T)$, approximately balances diabatic heating, H , to reach low values¹⁷, approaching the radiative equilibrium temperature T_e (refs 18, 19), with $H - C(T_e) = 0$. The situation changes abruptly when the Sun returns. Absorption of solar radiation by O_3 causes H to increase rapidly as the number of sunlit hours increases. However, radiative cooling to space, which is temperature-sensitive, remains close to the low value achieved during the polar night. The temperature evolution is constrained by the prognostic zonal momentum and energy equations through the thermal wind relationship and mass continuity equation. Thus, local temperature can adjust

only on a longer dynamical timescale, due to 'dynamical inertia'^{18,20}. In the absence of wave transports, the dynamical timescale is found²⁰ to be related to the newtonian cooling timescale of about 1–2 months for the Antarctic lower stratosphere²¹. The radiative imbalance temporarily leads to a positive net heating,

$$Q = H - C \approx \Delta H > 0 \quad (1)$$

where ΔH is the incremental increase in heating due to absorption of solar radiation.

The above increase in net heating may be balanced by an increase in local air temperature (T), or by a vertical ascent of the heated air mass or both. This follows from the zonal-mean thermodynamics equation, which is, neglecting wave transports,

$$\bar{Q} = \frac{\partial}{\partial t} \bar{T} + \Gamma \bar{w} \quad (2)$$

where $\Gamma = (\partial \bar{T} / \partial z) + 9.8 \text{ } ^\circ\text{C km}^{-1}$ is the static stability parameter, the overbar denotes the zonal mean and $\bar{w}$ is the residual mean vertical velocity. Because of the 'dynamical inertia', the net heating should lead mainly to an upward circulation with vertical velocity

$$\bar{w} \approx \bar{Q} / \Gamma \quad (3)$$

during early spring. The upward branch of this reverse circulation in the lower stratosphere occurs in the vortex where $\bar{T} < T_e$, while the downward branch occurs at the edge of the polar vortex where $\bar{T} > T_e$. As upwelling (subsidence) lowers (raises) the column density of O_3 , it follows from the above arguments that, during the Antarctic spring, cold (warm) temperature should be correlated with O_3 minimum (maximum) at the altitude where O_3 concentration is largest. This is consistent with observation³. In the northern polar stratosphere, winter temperatures are maintained by dynamical transport, leading to values $\sim 20\text{--}30 \text{ } ^\circ\text{C}$ warmer than the temperature that prevails over the southern winter pole. Radiative cooling thus tends to exceed heating in Arctic winter, resulting in a downward diabatic circulation, as evidenced by the movement of stratospheric aerosols²².

To estimate the effect of upwelling on O_3 , we use a simple model based on the zonal-mean equation for the O_3 mixing ratio $\bar{f}$,

$$\frac{\partial \bar{f}}{\partial t} + \bar{v} \frac{\partial \bar{f}}{\partial y} + \bar{w} \frac{\partial \bar{f}}{\partial z} = 0 \quad (4)$$

where the effect of irreversible mixing and photochemical reactions are neglected. As the residual mean meridional velocity ($\bar{v}$) is small near the pole, the effect of horizontal transport is also neglected. Figure 1 shows the calculated heating rate, ΔH , due to the absorption of solar radiation by ozone alone. The vertical advective velocity is calculated using equations (3) and (1), as described in Fig. 1 legend. Figure 2a, b shows a comparison of our calculated column densities of O_3 with observations at Syowa Station for 1982 (ref. 1) and with those at the latitude of Halley Bay for 1983 (ref. 3). The calculated rate of decline in September and October appears to be consistent with the observed zonal mean values³, but is smaller than the minimum over all longitudes.

An assessment of the photochemical contribution is hampered by the lack of kinetic data. The following discussion should be viewed as an assessment of the requirements on photochemistry for it to explain the observed O_3 behaviour. McElroy *et al.*⁴ have proposed that catalysis of O_3 recombination involving the bromine-chlorine cycle²³ may be important in Antarctic early spring:

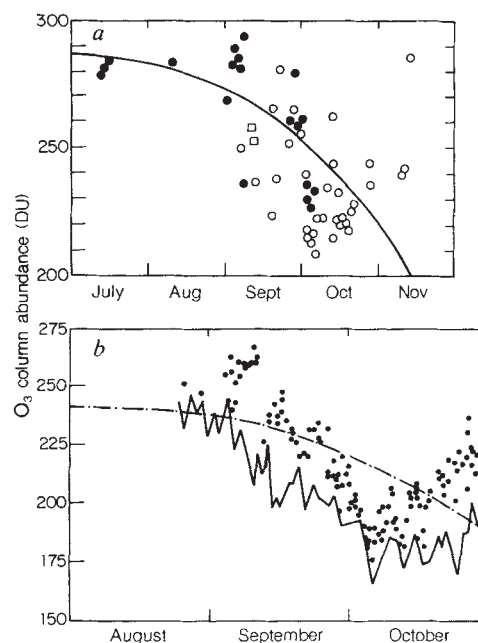
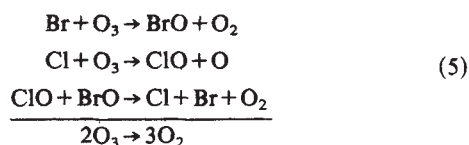


Fig. 2 a, Calculated column abundance (in Dobson units; 1 DU = 10^{-3} atm cm) of O_3 at 69° S (curve), compared with observations for the Syowa Station¹. The O_3 profile is calculated as described in Fig. 1 legend. The O_3 advected out of the top of the vortex (at ~ 30 km) is assumed to be swept away by horizontal motion and will not contribute to the column content. The data are from Dobson instrument by direct Sun ($\circ$), by cloudy zenith ($\square$) and by Moon ($\bullet$). b, Calculated column abundance of O_3 at 76° S (---) compared with observations from the Total Ozone Mapping Spectrometer (TOMS)³. The computation procedure is as stated in a, except that solar insolation is calculated for 76° S and the initial O_3 profile is adjusted to give an initial value of 240 DU for the column content. The data from TOMS³ comprise the minimum observed values for all longitudes between 74° S and 78° S (solid line) and observations at Halley Bay (dots).

In order for reaction (5) to provide a decrease of 0.5% per day in the column abundance of O_3 , the 24-h-averaged abundances of ClO and BrO at ~ 18 km must be ~ 0.4 p.p.b.v. and 20 p.p.t.v. (parts per 10^{12} by volume), respectively.

Results of our two-dimensional model¹¹ indicate a total chlorine (ClY) mixing ratio of 2.6 p.p.b.v. in the upper stratosphere and 1.5 p.p.b.v. at 18 km, of which 30% is in the form of ClO_x ($\text{ClNO}_3 + \text{Cl} + \text{ClO} + \text{OCIO} + 2 \times \text{Cl}_2 + \text{HOCl}$). If the observed decrease in O_3 were to be explained by photochemistry alone, ClO would have to be the major ClO_x species in early spring. The daytime abundances of the species Cl, ClO and ClNO_3 are primarily determined by the equilibrium between production of ClO by photolysis of ClNO_3 and rapid re-formation of ClNO_3 , with equilibrium achieved within 1–2 h of daylight²⁴. To maintain the high ClO concentration, we must consider processes which both convert ClNO_3 to ClO_y ($\text{Cl} + \text{ClO} + \text{OCIO} + 2 \times \text{Cl}_2 + \text{HOCl}$) and maintain sufficiently low NO_x to inhibit re-formation of ClNO_3 . It is also important to consider the initial conditions of chlorine and nitrogen species in late autumn when the polar vortex is established before entering polar night.

Figure 3 summarizes the possible chemical mechanisms controlling chlorine and nitrogen species from Antarctic late autumn to early spring. Net conversion of NO_x to N_2O_5 occurs in the late autumn due to the rapidly decreasing photolysis rates of N_2O_5 with increasing solar zenith angles and shorter days. As the NO_x decreases, the daytime ClO increases, until the ClO and NO_x concentrations are comparable. From that point on, all ClO and NO_x will be sequestered in ClNO_3 within 1–2 h of darkness. Under these conditions, we calculate daytime concentrations of ClO of 0.07 p.p.b.v. at 18 km during late autumn

11. Ko, M. K. W. Tung, K. K., Weisenstein, D. K. & Sze, N. D. *J. geophys. Res.* **90**, 2313-2329 (1985).
12. Solomon, S., Garcia, R. R. & Stordal, F. *J. geophys. Res.* **90**, 12981-12990 (1985).
13. Dütsch, H. U. *Pure appl. Geophys.* **116**, 511-529 (1978).
14. Bowman, K. P. & Krueger, A. J. *J. geophys. Res.* **90**, 7967-7976 (1985).
15. McIntyre, M. E. *J. met. Soc. Japan* **60**, 37-65 (1982).
16. McIntyre, M. E. & Palmer, T. N. *Nature* **320**, 593 (1983).
17. Murgatroyd, R. J. in *The Global Circulation of the Atmosphere* (ed. Corky, G. A.) (Royal Meteorological Society, London, 1969).
18. Fels, S. B., Mahlman, J. D., Schwarzkopf, M. D. & Sinclair, R. W. *J. atmos. Sci.* **37**, 2265-2297 (1980).
19. Wehrbein, W. M. & Leovy, C. B. *J. atmos. Sci.* **39**, 1532-1544 (1982).
20. WMO/NASA, *Atmospheric Ozone: Assessment of our Understanding of the Process Controlling its Present Distribution and Changes* (in the press).
21. Kiehl, J. T. & Solomon, S. *J. atmos. Sci.* (in the press).
22. Kent, G. S., Trepte, C. R., Farrukh, U. O. & McCormick, M. P. *J. atmos. Sci.* **42**, 1536-1551 (1985).
23. Yung, Y. L., Pinto, J. P., Watson, R. T. & Sanders, S. P. *J. atmos. Sci.* **37**, 339-353 (1980).
24. Ko, M. K. W. & Sze, N. D. *J. geophys. Res.* **89**, 11619-11632 (1984).
25. Clyne, M. A. A. & Watson, R. T. *JCS Faraday Trans. 1*, **73**, 1169 (1977).
26. Wofsy, S. C. *J. geophys. Res.* **83**, 364-378 (1978).
27. Turco, R., Hamill, P., Toon, O. B., Whitten, R. C. & Kiang, S. S. *J. atmos. Sci.* **36**, 699 (1979).
28. McCormick, M. P., Steele, H. M., Hamill, P., Chu, W. P. & Swisler, T. J. *J. atmos. Sci.* **39**, 1387-1397 (1982).
29. Hoffman, D. J. & Rosen, J. M. *Geophys. Res. Lett.* **12**, 13-16 (1985).
30. Mugnai, A., Fiocco, G. & Grams, G. Q. *Jl R. met. Soc.* **104**, 783-796 (1978).
31. Strobel, D. F. *J. geophys. Res.* **83**, 6225-6230 (1978).
32. Rowland, F. S., Sato, H., Ichwaja, H., Elliott, S. M. *J. Phys. Chem.* **90**, 1985-1988 (1986).

High-resolution lattice imaging reveals a 'phase transition' in Cu/NiPd multilayers

C. S. Baxter & W. M. Stobbs

Department of Metallurgy and Materials Science, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK

Considerable interest has been generated in metal multilayers by the discovery of anomalous variations in their physical properties as a function of layer thickness. Of particular interest is the 'supermodulus effect', observed for several multilayer systems at modulation wavelengths of 1.6-3.0 nm, in which an elastic modulus can increase by up to 400% (ref. 1). Various explanations for this effect have been suggested, and as part of an investigation aimed at clarifying the role of coherency strains in multilayer properties we are studying the behaviour of Cu/NiPd multilayers. We have developed transmission electron microscopy (TEM) techniques which allow the structural characterization of these materials, including the measurement of both the extent of interlayer mixing and the lattice plane spacings. Here we present some of our results in the latter category and discuss data which indicate that a novel form of 'phase transition' occurs as the modulation wavelength is decreased.

Metal multilayers typically consist of ~1,000 alternating layers of two metals or alloys, with modulation wavelengths (defined as the combined thickness of adjacent layers of each material) in the range 1.0-6.0 nm. Explanations which have been put forward for wavelength-dependent multilayer property changes range from theories based on critical Fermi surface/Brillouin zone interactions to arguments that consider the consequences of interlayer stress relaxation by the generation of misfit dislocations. The development of Fermi surface theories is hindered by the fact that the precise form of multilayer Fermi surfaces is not known, but as a first approximation these can be taken as equivalent to those of the corresponding homogeneous alloys. For example, it has been shown² that the peaks in modulus of both Ag/Pd and Cu/Ni multilayers occur at modulation wavelengths for which the additional Brillouin zones introduced by the modulations make tangential contact with the alloy Fermi surfaces. Any theoretical approach necessitates the selection of appropriate structural models to represent the multilayers, so we have had to develop new TEM techniques allowing better structural characterization of a multilayer than is possible by X-ray methods. The Cu/NiPd multilayer system has the added advantage in this context that the difference

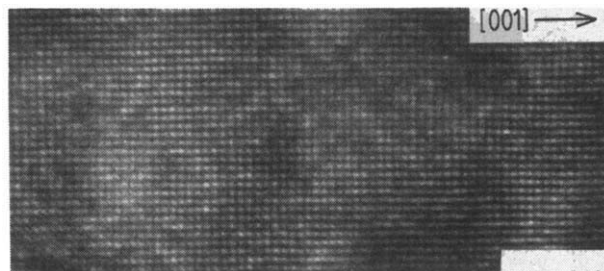


Fig. 1 Non-axial lattice image of a Cu/NiPd multilayer taken on the Cambridge HREM operated at 500 kV. Scale bar, 2 nm.

between the bulk lattice parameters of the layers (and hence the resulting coherency strains) can be altered systematically by varying the Ni:Pd ratio. Our observations indicate that interfacial dislocations do not play a critical role in the behaviour of the Cu/NiPd system at or near critical wavelengths, and this result is discussed elsewhere³.

The multilayers we have used were deposited by Somekh using a magnetron sputter deposition system⁴. Although we have examined films grown with both [001] and [111] epitaxial orientations, the results we report here are for [001]-oriented films only.

A fundamental part of our structural studies has been an attempt to locate the atomic positions as accurately as possible using high-resolution electron microscopy (HREM). Obtaining analysable high-resolution images of a modulated structure, even when thin enough to be treated as a weak-phase object, requires better resolution than is needed for the measurement of lattice spacings in either layer alone. This is because the modulated spacing variations are associated primarily with the high-frequency Fourier coefficients. Simulations for idealized specimens (that is, those exhibiting no variations in phase caused by, for example, local thickness changes) indicate that the lattice plane spacing variations should be detectable in axial lattice images, although the fringe positions would be too sensitive a function of defocus to be analysed reliably. In fact, despite sufficient resolution for accurate imaging of the spacings concerned in uniform specimens, in practice the variation cannot be detected⁵—presumably because when working near the limit of the transfer function the sensitivity of the image to inevitable local phase changes is accentuated. We have overcome this problem by using non-axial rather than axial illumination conditions, which increases the accuracy to which lattice spacings can be measured⁶; a typical image of a 4.4-nm modulation-wavelength Cu/NiPd multilayer is shown in Fig. 1. Images of this type now contain modulated fringe spacings, as revealed by the measurements shown in Fig. 2b, obtained from the fringe intensity profile shown in Fig. 2a. Our analysis of these profiles has allowed us to determine the lattice strains of multilayers with relatively large modulation wavelengths and lattice mismatch⁷, but for specimens for which either the lattice mismatch is small or, more frustratingly, the wavelength is low (including the anomalous range near 2.0 nm), the modulation cannot be detected above the random noise in the fringe measurements. However, in these cases it is still possible to gain significant structural information by measuring the average lattice spacings in directions parallel and perpendicular to the layer normal ($\bar{d}(002)$ and $\bar{d}(020)$ respectively). The lattice mismatch between the layers in our multilayers is accommodated by coherency strains, and conventional elasticity theory can be used as a first approximation to predict the expected values of these average spacings, which will not generally be the same. The relative values of $\bar{d}(002)$ and $\bar{d}(020)$ can be measured from lattice images with a high degree of accuracy ($\sim \pm 0.2\%$), and when this is done for a relatively long-wavelength (4.4 nm) Cu/NiPd multilayer we find that the measured difference in the two spacings has the same sense and size ($\bar{d}(020)$ 1% bigger than $\bar{d}(002)$) as

Fig. 2 *a*, Intensity profile of lattice fringes from a Cu/NiPd multilayer as used for measuring fringe positions. *b*, Deviations from line of least-squares fit of fringe positions measured from *a*; the mean fringe spacing is 6.6 pixels. To show the modulation clearly, these data have been presented as the deviation of each successive fringe position from the line of least-squares fit through the fringe positions: in this format, the maximum spacing is associated with the maximum positive gradient and the minimum spacing with the maximum negative gradient. As with any high-resolution imaging, care is needed in the interpretation of these measurements and the measured fringe spacings cannot be directly related to the lattice spacings in the multilayer. Instead, the spacings measured from a focal series of images were compared with spacings in images simulated using multilayer models with different interlayer strains. These models were based on composition profiles determined by a Fresnel method^{5,7}. For the image series including the example shown in Fig. 2, the best match between the simulated and measured spacing variations (9% in Fig. 2*b*) was found for a model with a 6% variation in lattice plane spacing. The fit with a value of 6% is less than would have been expected for a nominal misfit of 3.5% (as for Cu/Ni_{0.5}Pd_{0.5}), but the reduced value is presumably explained by surface relaxation in the TEM specimen⁸.

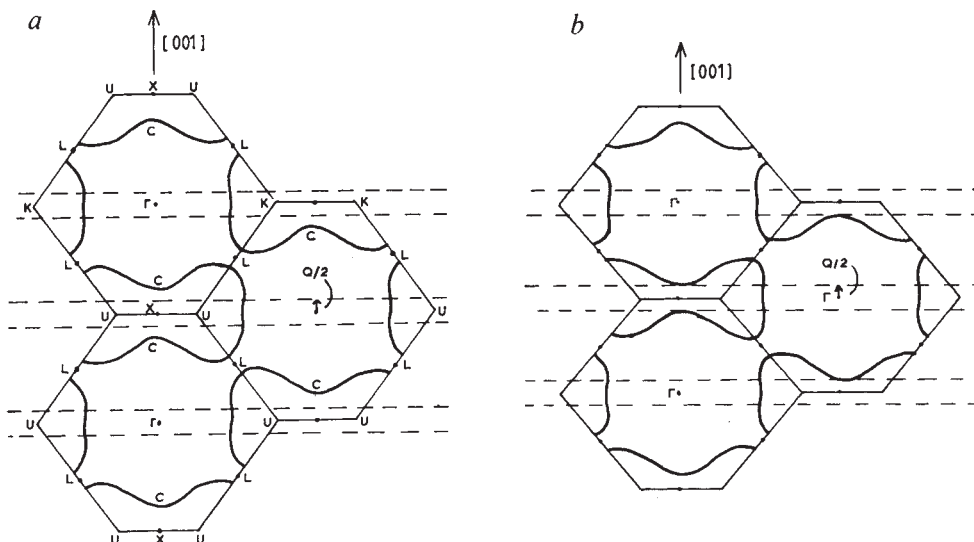
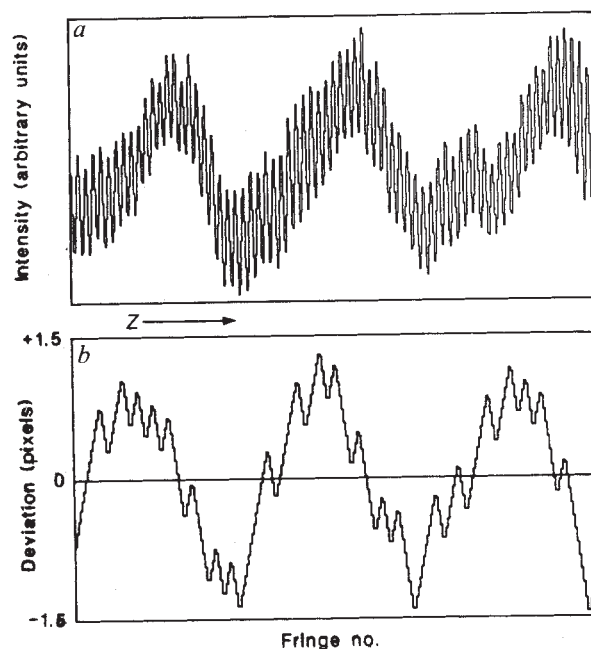


Fig. 3 *a*, Possible Fermi surface of Cu/NiPd drawn in the periodic zone scheme (based on Fermi surface of CuNi (ref. 9) and with Γ , X, L, K and U used conventionally to indicate high-symmetry points of the Brillouin zone). The zone boundary introduced by a 1.6-nm wavelength modulation along [001] (wave-vector Q) are also shown. No points of tangential contact occur, but contact could occur at C with a shorter-wavelength modulation. *b*, Possible Fermi surface of Cu/NiPd with a tetragonal distortion of the unit cell such that the mean (002) lattice spacing is increased by 10%. The zone boundaries corresponding to the 1.6-nm modulation now make contact with the Fermi surface.

predicted by our linear elasticity calculations. However, at the shorter wavelength of 1.6 nm, $\bar{d}(002)$ is 2.5% greater than $\bar{d}(020)$, and this cannot be accounted for using conventional elasticity theory. We are convinced that this result reflects a genuine wavelength-dependent change in the structure of the specimens rather than an electron-optical imaging artefact, because both differences were found to be consistently measurable to an accuracy far greater than the random errors involved for a range of specimen thickness, objective defocus and illumination conditions. Furthermore, surface relaxations, although undoubtedly important⁸, could not provide an explanation for this phenomenon: such effects are most marked when the specimen thickness is comparable to the wavelength, but the anomalous spacings are found in short-wavelength multilayers where the specimen thickness is at least 10 times the wavelength. We therefore conclude that an anomalous change has occurred in the structure of the shorter-wavelength multilayer, which can be thought of as a transformation to a novel tetragonal 'phase'. The transformation is difficult to model because of uncertainties in the electronic structure, but a consistent qualitative interpretation can be obtained by examining the consequences of the observed structural distortion on possible Fermi surface/Brillouin zone interactions. As a starting point, we have taken the Fermi surface of a Cu/NiPd multilayer as being approximately

equivalent to that of a homogeneous CuNi alloy: this is shown in Fig. 3*a* for an undistorted cubic unit cell, together with the Brillouin zone introduced by a 1.6-nm modulation along [001]. No points of critical contact occur at this wavelength, but contact would occur at the points marked 'C' for shorter-wavelength (longer-wave-vector) modulations. However, critical contact is obtained at a wavelength of 1.6 nm if a tetragonal distortion of the crystal structure is introduced such that the dimensions of the Fermi surface remain approximately constant with respect to the origin but the lattice dimension along [001] is increased by ~10% (see Fig. 3*b*). The distortion required for contact is more than twice the observed change in the mean spacings (~4%), but both the substitution of Pd for Ni and the presence of the lower band gaps introduced by the modulation will raise the Fermi energy, while the Fermi surface would be expected to be increasingly distorted towards the Brillouin zone boundary as the modulation wavelength is decreased towards the value at which critical contact can occur by a 'transformation' of structure. This argument is clearly grossly oversimplified and a full calculation of any accuracy would be very difficult, but it is apparent that the trends indicated are in the observed direction.

This picture has several consequences: for example, while the transformation would be first-order thermodynamically, given

the associated changes in volume, precursory effects would be expected which should be reflected in progressive changes in physical properties on either side of the critical wavelength (as are, in fact, generally observed). A transformation-induced mechanism for the changes in physical properties of otherwise isostructural metal multilayers could thus have general significance. It is not clear whether the transformation we have observed occurs in one or both components of the multilayer, but it might be possible to design a system for which both layers exhibit transformations at different wavelengths. Any correlated anomalies in the physical properties, if observed, would be useful in testing the ideas outlined.

Whatever the precise mechanism for the 'transformation', our data demonstrate a fundamental change in the structure of Cu/NiPd multilayers as a function of the modulation wavelength. Experiments are in hand to correlate the wavelength dependence of the structural and physical anomalies of this system.

Received 8 April; accepted 25 June 1986.

1. Henein, G. E. & Hilliard, J. E. *J. appl. Phys.* **54**, 728-733 (1983).
2. Pickett, W. E., *J. Phys.* **F12**, 2195-2204 (1982).
3. Baxter, C. S. & Stobbs, W. M. *Pap. 44th A. Meet. Electron Microsc. Soc. Am.* (San Francisco Press).
4. Somekh, R. E. & Baxter, C. S. *J. Cryst. Growth* (submitted).
5. Baxter, C. S. & Stobbs, W. M. *EMAG 85*, 387-390 (Institute of Physics Conf. Ser. No. 78, 1985).
6. Hall, D. J., Self, P. G. & Stobbs, W. M. *J. Microsc.* **130**, 215-224 (1983).
7. Baxter, C. S. Thesis, Univ. Cambridge (1986).
8. Treacy, M. M. J., Gibson, J. M. & Howie, A. *Phil. Mag.* **A51**, 389-417 (1985).
9. Gordon, B. E. A., Temmerman, W. M. & Gyorffy, B. L. *J. Phys.* **F11**, 821-857 (1981).

New method for the measurement of osmium isotopes applied to a New Zealand Cretaceous/Tertiary boundary shale

Frederick E. Lichte*, Shane M. Wilson†, Robert R. Brooks†, Roger D. Reeves†, Jiri Holzbecher‡ & Douglas E. Ryan‡

* US Geological Survey, Federal Center, Denver, Colorado 80225, USA

† Department of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand

‡ Trace Analysis Research Centre, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1

The determination of osmium content and isotopic abundances in geological materials has received increasing attention in recent years following the proposal of Alvarez *et al.*¹ that mass extinctions at the end of the Cretaceous period were caused by the impact of a large (~10 km) meteorite which left anomalously high iridium levels as a geochemical signature in the boundary shales. Here we report a new and simple method for measuring osmium in geological materials, involving fusion of the sample with sodium peroxide, distillation of the osmium as the tetroxide using perchloric acid, extraction into chloroform, and absorption of the chloroform extract onto graphite powder before instrumental neutron activation analysis. In a variant of this technique, the chloroform extract is back-extracted into an aqueous phase and the osmium isotopes are determined by plasma-source mass spectrometry (ICPMS). We have used this method on the Woodside Creek (New Zealand) Cretaceous/Tertiary boundary clay and have obtained the first osmium content (60 ng g⁻¹) for this material. The ¹⁸⁷Os/¹⁸⁶Os ratio is 1.12 ± 0.16, showing a typical non-crustal signature. This combined distillation-extraction-ICPMS method will prove to be useful for measuring osmium isotopes in other geological materials.

Although ~50 Cretaceous/Tertiary (K/T) boundary sites with anomalous iridium levels have now been discovered², only one of these, at Woodside Creek, New Zealand, is on land in the Southern Hemisphere and is therefore of some importance for an interpretation of the events which occurred at the end of the Cretaceous. The geochemistry of the Woodside Creek occurrence has been reported by Brooks *et al.*³, who found 153 ng g⁻¹ iridium (carbonate-free basis) in the basal layer of the 8-mm-thick boundary shale, together with an integrated iridium deposition of 187 ng cm⁻².

Most of the original work on the K/T boundary in different parts of the world involved the use of iridium as a marker, particularly as this element is easier to determine by neutron activation analysis (NAA) than are the other platinum metals. Osmium, which has a similar abundance to iridium, serves almost equally well as an indicator of extraterrestrial (or mantle) material, but has a sensitivity ~30 times poorer than that of iridium when determined by neutron activation. The isotopic ratio ¹⁸⁷Os/¹⁸⁶Os is also of use in explaining the nature of geochemical signatures, as meteorites (and mantle material) typically have ratios near 1.0 (ref. 4), whereas crustal material averages ~10.0 (refs 5, 6).

The recent development of ICPMS for determining isotopic abundances⁷ has opened up exciting new possibilities for determining these isotopes in geological materials. We have therefore developed a technique of removing osmium from rocks by distillation, followed by ICPMS or instrumental NAA (INAA) determinations. This procedure has been applied to the measurement of osmium isotopes in the Woodside Creek K/T boundary shale, as described below.

Samples of the basal 2-mm layer of the boundary shale (0.5 g) were fused with 5 g of sodium peroxide at 650 °C for 2 min. After cooling, the melt was transferred to a 50-ml flask together with 20 ml of 70% perchloric acid, and distillation was carried out into a receiver initially containing 2 ml of the same acid. After 8 ml of mixture had been recovered, distillation was terminated and the distillate was shaken for 5 min with 10 ml of chloroform. The extract was used either for INAA or for ICPMS determinations. For the former purpose the extract was mixed with 10 ml of a 1:1 chloroform/ethanol reagent containing 0.5% thiourea acidified with perchloric acid (2% v/v), and the combined mixture was absorbed and evaporated drop by drop on 1 g of high-purity graphite powder. The powder was analysed for osmium by INAA using a Slowpoke reactor with a flux of 5 × 10¹¹ n cm⁻² s⁻¹. The irradiation time was 16 h and counting was performed for 16 h or longer using a Ge(Li) detector or LEPD (low-energy particle detector) after a decay period of 7 days.

For the ICPMS determinations the chloroform extract was back-extracted into 2 ml of an aqueous phase containing 1% thiourea and 0.5 M sulphuric acid. The spectrometer was a Sciex Elan ICP-MS instrument with a glass frit nebulizer⁸ which

Table 1 Osmium, iridium and gold concentrations (ng g⁻¹) and osmium isotopic ratios in meteorites and Cretaceous/Tertiary boundary shales

	Woodside Creek†	Allende chondrite	Stevens Klint	All chondrites
Au	59 ^{3‡}	137 ⁹ , 145 ¹⁰	7 ¹¹	—
Ir	127 ^{3‡}	745 ¹² , 785 ¹⁰	46 ¹¹	—
Os	60*	720*, 759 ¹² , 828 ¹⁰	60 ¹¹	—
Os/Au	1.12*	4.96*	8.57 ¹¹	—
Ir/Au	2.15 ³	5.13 ^{7,10}	6.57 ¹¹	—
¹⁸⁷ Os/ ¹⁸⁶ Os	1.12*	0.91*	1.65 ⁵	1.11 ⁴

Superscript numbers are references.

* This work.

† Woodside Creek CaCO₃ is 17%, not 7% as quoted in ref. 3.

‡ Values from which the CaCO₃-free concentrations of ref. 3 were derived.

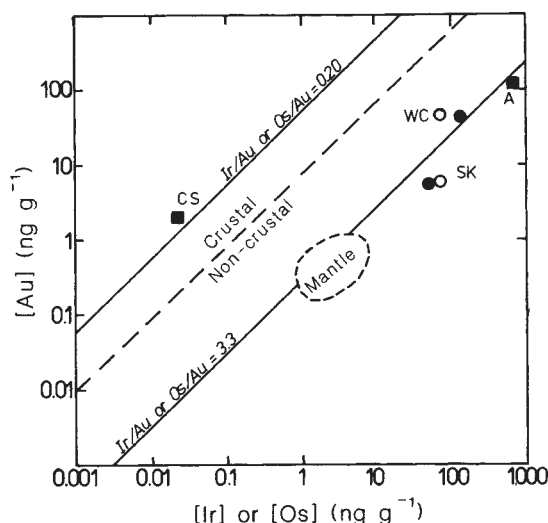


Fig. 1 Plots of gold versus iridium (●) or osmium (○) in crustal, meteoritic and K/T boundary material (■, both Ir and Os). CS, mean for Canadian Shield⁷; SK, Stevens Klint⁸; WC, Woodside Creek (this work); A, Allende carbonaceous chondrite (this work).

increased nebulization efficiency to 90%, compared with 1–3% for conventional nebulizers. Each sample was analysed 10 times with a total integration time of 15 min at a nebulization rate of 30 $\mu\text{l min}^{-1}$. Standard deviations (2σ) were calculated automatically by the on-board computer. Other instrumental conditions were: r.f. power 1,150 W, nebulizer gas flow rate 0.7 litre min^{-1} , plasma gas 13 litre min^{-1} , sheath gas 1.0 litre min^{-1} , plasma sampling distance 22 mm from the load coil. The instrument was calibrated using an osmium standard dissolved in thiourea and sulphuric acid to match the sample solutions. The normalized instrument response was verified against the Allende meteorite standard, which has documented isotopic abundances. A reagent blank showed no detectable blank for any osmium isotope.

Table 1 lists the data for total iridium and osmium and for the $^{187}\text{Os}/^{186}\text{Os}$ isotopic ratios. The osmium data are the first reported for the New Zealand material. Our data for the Allende carbonaceous chondrite and other data for K/T boundary material from Stevens Klint, Denmark, are also given.

Experiments with samples spiked with known amounts of osmium indicated a recovery of 93% in the distillation-extraction procedure. After allowing for this factor, we obtain a total osmium content of 720 ng g^{-1} (by INAA) for the Allende specimen, which is close to the values of 745 ng g^{-1} (ref. 9) and 828 ng g^{-1} (ref. 10) reported by other workers. Our value of 60 ng g^{-1} for osmium in the Woodside Creek basal layer is identical to that found by Kyte *et al.*¹¹ for Stevens Klint, and is in line with the observation of Brooks *et al.*³ that the New Zealand and Danish K/T shales have similar chemical compositions (apart from the higher CaCO_3 content of the latter).

Palme *et al.*¹² have pointed out that the Ir/Au ratio in geological materials is a reliable geochemical signature by which to separate crustal and non-crustal samples. As the osmium and iridium concentrations in chondrites are almost identical ($\sim 700 \text{ ng g}^{-1}$), the Os/Au ratio should be an equally reliable marker for non-crustal materials. Figure 1 shows a plot of gold versus iridium or osmium for the New Zealand and Danish boundary shales, as well as for the Allende meteorite and for crustal material represented by mean values for the Canadian Shield¹². The data clearly indicate a non-crustal source for the New Zealand material, and because the values lie well outside the field for mantle material, an extraterrestrial source is indicated.

Our value of 1.12 ± 0.16 (2σ) for the $^{187}\text{Os}/^{186}\text{Os}$ isotopic ratio is the same as that reported by Luck and Allegre⁶ for the mean

of 10 chondrites (1.11 ± 0.01), suggesting an extraterrestrial source for the osmium. It is true that the ratio is also typical of mantle material but this question has already been dealt with above (see also Fig. 1) and we prefer to propose a probable extraterrestrial origin for the osmium. Certainly the findings clearly negate any arguments that the osmium is of crustal origin. Using our value of 10.4% for the extraterrestrial component of the New Zealand boundary shale³, we can calculate that correction for terrigenous osmium would lower our ratio from 1.12 to 1.10. This is still extremely close to the ratio for chondrites and gives the best evidence so far for the non-crustal origin of the osmium (and by inference also probably the iridium) in the Woodside Creek material.

It is suggested that our combined distillation-extraction-ICPMS procedure will have a continuing use for the determination of osmium isotopic ratios in a wide variety of geological materials. In the application reported here, measurements were made on solutions containing only $\sim 0.15 \text{ ng ml}^{-1}$ of the two osmium isotopes (that is, $\sim 5 \text{ ng}$ total osmium), and our combined procedure should prove to be an acceptable alternative to the chemical separation-mass spectrometric ion probe method now in use.

Received 31 December 1985; accepted 30 May 1986.

1. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
2. Alvarez, W., Alvarez, L. W., Asaro, F. & Michel, H. V. *Spec. Pap. geol. Soc. Am.* **190**, 305–315 (1982).
3. Brooks, R. R. *et al.* *Science* **226**, 539–542 (1984).
4. Luck, J. M. & Allegre, C. J. *Nature* **302**, 130–132 (1983).
5. Luck, J. M. & Turekian, K. K. *Science* **222**, 613–615 (1983).
6. Palmer, M. R. & Turekian, K. K. *Nature* **319**, 216–220 (1986).
7. Date, A. R. & Gray, A. L. *Analyt. Lond.* **108**, 159–165 (1983).
8. Layman, L. R. & Lichte, F. E. *Analyt. Chem.* **54**, 638–642 (1982).
9. Takahashi, H., Janssens, M. J., Morgan, J. W. & Anders, E. *Geochim. cosmochim. Acta* **42**, 97–106 (1978).
10. Kallemeyn, G. W. & Wasson, J. T. *Geochim. cosmochim. Acta* **45**, 1217–1230 (1981).
11. Kyte, F. T., Zhou, Z. & Wasson, J. T. *Nature* **288**, 651–656 (1980).
12. Palme, H., Janssens, M. J., Takahashi, H., Anders, E. & Hertogen, J. *Geochim. cosmochim. Acta* **42**, 313–325 (1978).

Silicate microspherules intercepted in the plume of Etna volcano

R. Lefèvre*, A. Gaudichet*† & M. A. Billon-Galland†

* Laboratoire de Microscopie Analytique Appliquée aux Sciences de la Terre, Université Paris XII, 94010 Créteil, France

† Laboratoire d'Etude des Particules Inhalées, 44, rue Charles Mourau, 75013 Paris, France

A possible volcanic origin has been suggested for the micrometre-sized spherules that have been discovered in polar snows or ice-cores^{1–5}, and in the stratosphere^{6,7}. However, although such particles, especially when black and magnetic, have been identified among the tephra deposits of some volcanoes^{3,8–12}, their volcanic emission has never been directly observed. Here we describe microspherules intercepted in the plume of Mount Etna, Sicily, during moderate recurring volcanic activity. Their study, using analytical transmission and scanning electron microscopy (ATEM and ASEM) demonstrates the simultaneous presence in the plume of glassy silicate microspherules of various chemical compositions (47–98% SiO_2).

Particles were collected in September 1984 and August 1985 on a Nuclepore membrane (porosity 0.4 μm) by bulk filtration of the aerosol flux emitted by a summit crater of Mount Etna (Bocca Nuova, elevation 3,340 m above sea level). A volume of 2–3 m^3 of diluted volcanic plume was filtered at a flow rate of 1 $\text{m}^3 \text{ h}^{-1}$. The sampling head was placed inside the crater at the end of a 6-m-long flexible fishing rod, 2 m beneath and 2 m away from the crater rim. Because the wind blowing onto the crater was weak or moderate during sampling, only rising

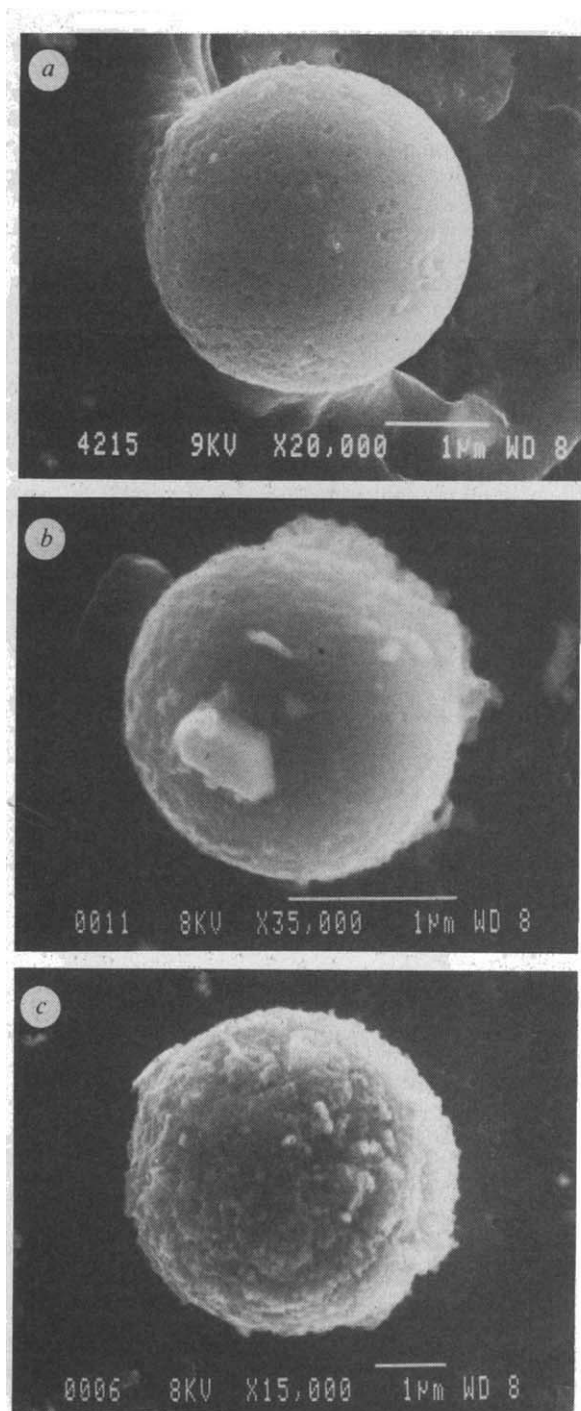


Fig. 1 Scanning electron micrographs of three typical microspherules intercepted in the permanent plume of Mount Etna volcano and washed with filtered distilled water (JEOL JSM 840 electron microscope; accelerating voltage, 8 kV; scale bar, 1 μ m). *a*, Smooth and clean microspherule showing small punctuations; *b*, Microspherule coated with microparticles, among them a probable feldspar crystal. *c*, Granular microspherule.

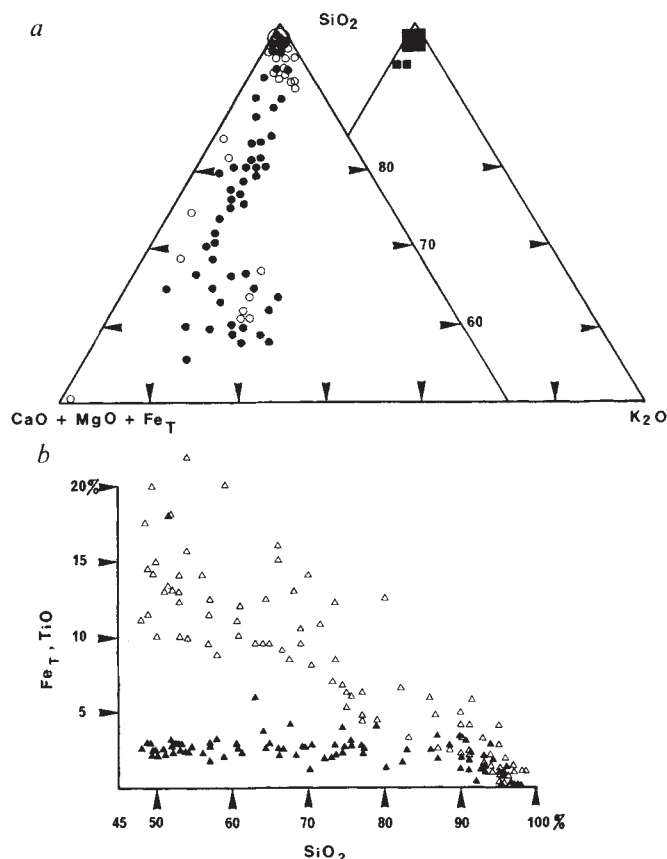


Fig. 2 *a*, Chemical compositions plotted in an SiO_2 versus K_2O and $\text{CaO} + \text{MgO} + \text{Fe}_T$ diagram of 54 glassy microspherules intercepted in September 1984 ($\bullet$) and 36 in August 1985 ($\circ$), compared with 17 shards of volcanic glass intercepted in September 1984 ($\blacksquare$), in the plume of Mount Etna. These diagrams show the variation of the compositions of microspherules and the hyper-siliceous composition of shards. *b*, Relationship between SiO_2 content and Fe_T ($\blacktriangle$) or TiO ($\triangle$) of 90 microspherules. Note the Fe_T decrease linked with the SiO_2 increase. Titanium remains almost constant, except for the most SiO_2 -enriched spherules.

ation. For analysis by ASEM the membrane was not dissolved and was placed directly on the specimen holder.

Among the collected particles, non-silicate species (such as sulphates and oxides) are predominant. The silicates consist of crystals (such as feldspars), shards of silicate glass¹⁵ and microspherules. The diameter of the microspherules ranges from 0.1 to 5 μm and is $<2 \mu\text{m}$ for 85% of them. The concentration of microspherules was estimated to be $3 \times 10^6 \text{ m}^{-3}$ in September 1984 and $7 \times 10^6 \text{ m}^{-3}$ in August 1985. Their external morphology is either smooth and clean (Fig. 1*a*), coated with fine grains or crystals (Fig. 1*b*), or granular (Fig. 1*c*). Electron microdiffraction demonstrates the glassy structure for the smallest spherules (0.5 μm in diameter). The individual quantitative chemical composition of the microspherules was obtained by X-ray energy-dispersive spectrometry using the peak ratio method and taking Si as an internal reference element¹⁶. This is valid if the particles can be considered to be thin layers ($<2 \mu\text{m}$). The results are highly variable, even within a given sample. In particular, the SiO_2 content ranges from 47 to 98%, that is, from values typical of Etnan lavas (but with a noticeable K-enrichment) to values corresponding to nearly pure silica (Table 1 and Fig. 2). No correlation is observed between morphology, size and composition.

Assuming that each microspherule has a mean diameter of 1.4 μm and a mean density of 2.75 g cm^{-3} (between 2.85 for basalt and 2.26 for tridymite), the 10^6 microspherules we have

aerosols were collected and these showed no evidence of contamination by wind-blown tephra from the external flanks of the volcano.

The Nucleopore membrane, which has been covered with a carbon layer before sampling, received a second carbon layer after particle collection. The collected particles, trapped between the two carbon layers, were then transferred onto gold grids for analysis by ATEM^{13,14}. This transfer was performed by chloroform dissolution of the membrane under moderate aspir-

encountered per m³ of filtered air have a total mass $\sim 4 \times 10^{-3}$ mg. The SO₂ discharge from Mount Etna averages $\sim 4 \times 10^3$ tonne day⁻¹, with a mean SO₂ concentration in the plume^{17,18} of ~ 10 mg m⁻³. The discharge of microspherules can then be roughly estimated at 1.6 tonne day⁻¹.

The genesis of such microspherules and their surprisingly variable chemical composition within the same sample of diluted volcanic plume need to be explained. Some possible hypotheses are discussed below (Fig. 3):

(1) Contamination of the volcanic plume by industrial fly-ash seems to be excluded: the concentration of fly-ash samples in the same conditions at the periphery of the volcano and the limit of its sedimentary basement (Village of Maletto, elevation 975 m) is a $\sim 10^4$ m³, much lower than the concentration of glassy microspherules in the Mount Etna plume. Moreover, the chemical composition of industrial fly-ash is not commonly hypersiliceous^{19,20}, in contrast to a large proportion of our volcanic microspherules. A volcanic origin of these particles is therefore likely.

(2) The spheroid shape of the particles may be explained by two physical processes of emission: lava fountaining and/or phreato-magmatic eruptions. (a) Spheroidal droplets of glass are produced by highly fluid lava fountaining, as for example in Hawaii^{11,12,21,22}. Intense degassing of fluid lavas allows the dispersal of a spray of liquid magma droplets, which are then rapidly cooled to form simple glass spheres²³. The same mechanism has been proposed for the glass spheres discovered in Apollo 15 and 17 lunar samples²¹. However, the analysis of both terrestrial and lunar spheres has shown that their composition is constant. Therefore, although the mechanism of lava fountaining could be responsible for the presence of microspherules in the Etna plume, it does not explain the chemical variation. (b) Spheroidal-shaped particles have been observed²⁴ by SEM in the debris collected after experimental explosions, simulating phreato-magmatic eruptions produced by the reaction of water with a melted mixture of Al₂O₃ + Fe. These particles are very similar to hydrovolcanic ash²⁵. A phreato-magmatic process may thus explain the shape of our microspherules, but cannot account for their variable composition. (c) Whatever the genesis of microspherules (lava fountaining or phreato-magmatic explosions), the stagnation in the air-plume column between magma surface and crater rim of microspherules emitted by magmas of various compositions following one another in the pipe, is not compatible with the monotonous chemistry of the Etna basaltic lavas produced by lateral eruptions^{26,27} (Table 1). However, note that siliceous xenoliths derived from the sedimentary substratum of Mount Etna have been described in the pyroclastic

products of Mount Rossi and Mount Silvestri eruptions (1669 and 1892), on the southern flank of the volcano²⁸. The mechanism of generation and extraction of this material, in micro-spherical form, from such partially assimilated xenoliths is difficult to imagine.

(3) Incipient liquid immiscibility has also been suggested²⁹ to explain the presence of SiO₂-rich spherules (in the micrometre size range) found in the products of the experimental alteration of Etna basaltic glass by hot water in a Soxhlet apparatus. If such microspherical domains were present in the magma, they would be emitted in the plume directly after solidification at the surface. Such an explanation seems implausible in the present circumstances.

(4) Pre-existent microspherules could be extracted from weathered tephra deposits falling onto the magma surface before explosive or phreato-magmatic eruptions. The same microspherules found in the weathered tephra deposits on the inner flanks of the crater could be mobilized by the ascending gaseous flux which forms the permanent plume of the Mount Etna volcano. This hypothesis appears to be unlikely, as such structures have not been described in weathered volcanic tephra.

(5) Volcanic water emission by the Mount Etna plume has been evaluated¹⁸ at $\sim 1.4\text{--}2 \times 10^5$ tonne day⁻¹. This water is mostly meteoric in origin. During its circulation through the volcanic edifice, it would become heated and enriched in silica and ionic salts. At the magma/atmospheric interface, the water is released as vapour, and we may consider that dissolved silica-rich products could then be condensed as microspherules with varying chemical composition.

(6) Some experiments³⁰ emphasize the role of sulphur reacting with many of the common rock-forming silicates over wide temperature and pressure ranges to produce sulphides, oxides and pure silica. Depending on the amount of SO₂ discharge from Mount Etna^{17,18}, the sulphurization may be an important and suitable process for the generation of SiO₂-enriched microspherules and, depending on the speed of transport, may explain the observed compositional spectrum.

(7) Combustion of volcanic gases at the magma/atmosphere interface may be responsible for an increase in temperature, causing differential fusion of the solid silicates present in the magma, followed by a quenching of microspherules of various compositions. This mechanism is invoked to explain the fly-ash formation in coal-burning power-station furnaces: their morphology is acquired during cooling of the non-combustible mineral content of the coal, melted at temperatures $>1,500$ °C, while being carried along with the gaseous combustion products^{31,32}.

Table 1 Typical chemical compositions (oxide weight) of small volcanic glass shards (A to C) and microspherules intercepted in September 1984 (D to H) and August 1985 (I to M) in the permanent plume of Mount Etna, compared with the mean bulk composition of recent lavas and tephra (N)^{26,27}

	Shares of volcanic glass (September 1984)			Glassy microspherules (September 1984)					Glassy microspherules (August 1985)					Etna lavas
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
SiO ₂	93.8	93.3	95.2	96.3	82.7	71.6	64.5	49.3	95.3	86.6	70	60.6	49.8	47.8
Al ₂ O ₃	2.4	3.1	3	2.2	3.9	5.1	9.5	14.3	0.3	1.7	11.7	11.3	14.2	17.3
FeO _T	1.1	0.9	0.3	0	3.4	10.9	12.5	11.4	0.7	2.4	13.7	11.8	10.2	10.8
MgO	0	0	0	1	0.4	0	0.7	2.1	0	0	0	2.1	2.3	5.2
CaO	0	0.3	0.1	0.1	1.7	4	2.6	8.2	0.1	0.9	0.2	5	6.7	10.4
Na ₂ O	0.9	0.8	0	0	0	0	0	0	0	0	0	0	0.7	4
K ₂ O	1.5	1.4	1	0	5	5.7	7.2	11.8	2.5	5.2	3.1	6.5	13.3	1.7
TiO ₂	0	0	0.1	0	2.5	2.3	2.7	2.5	0.7	2.9	1.1	2.3	2.4	1.7
MnO	0	0	0	0.2	0	0	0	0	0	0	0	0.1	0	—
Total	99.7	99.8	99.7	99.8	99.6	99.6	99.7	99.6	99.6	99.7	99.8	99.7	99.6	99.1

The volcanic glass shards are always hypersiliceous; the microspherules vary from a typical Etna composition (with K-enrichment) to hypersiliceous composition (the Na-depletion is probably an artefact). Analyses were performed by ATEM (JEOL 100C fitted with an energy-dispersive x-ray spectrometer EDAX 711).

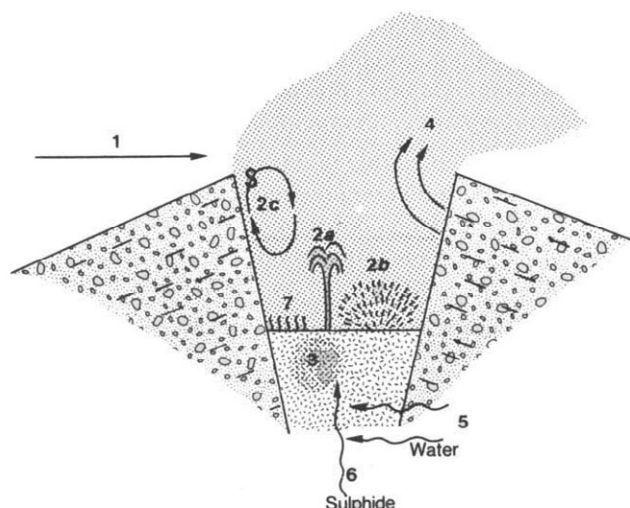


Fig. 3 Illustration of the hypotheses proposed in the text to explain the genesis and variable chemical composition of microspherules intercepted in the permanent plume of Mount Etna. (1), Contamination by industrial fly-ash; (2a) lava fountains; (2b) phreato-magmatic explosions; (2c) stagnation in the crater of different generations of microspherules emitted by successive magmas; (3) extraction of microspherules from SiO_2 -enriched domains in the magma (incipient liquid immiscibility); (4) extraction of pre-existent microspherules from weathered tephra; (5) dissolution by meteoric water, followed by water-emission and condensation of silica-rich products; (6) interaction between sulphur and silicates and/or magma to produce sulphides, oxides and silica; (7) combustion of volcanic gases and quenching of magmatic silicates; S, sampling location.

In the present state of our study, the more plausible hypotheses seem to be (2a), (5), (6) and (7) or a combination of these: the dissolving properties of water for silica at high temperatures and sulphurization of silicates are probably involved to produce SiO_2 -enriched microspherules; the lava fountains^{33,34} and similar phenomena involving gas emissions occur at Etna and, finally, the surprising morphological analogies between Etnean glassy microspherules and industrial fly-ash generated in power plant furnaces suggest similar thermal conditions.

This work was supported by the Centre National de la Recherche Scientifique, Paris: Associated Unity 717, Physico-Chemistry and Geochemistry of the Atmosphere and Programme Interdisciplinaire de Recherches sur la Prévision et la Surveillance des Eruptions Volcaniques. We thank Professor D. Velde, Dr P. Allard and Dr K. H. Wohletz for stimulating discussions.

Received 27 January, accepted 15 May 1986.

1. Fredriksson, K. & Martin, L. R. *Geochim. cosmochim. Acta* **27**, 245–248 (1963).
2. Wright, F. W. & Langway, C. C. *J. geophys. Res.* **68**, 19, 5575–5587 (1963).
3. Hodge, P. W. & Wright, F. W. *J. geophys. Res.* **69**, 12, 2449–2454 (1964).
4. Hodge, P. W., Wright, F. W. & Langway, C. C. *J. geophys. Res.* **69**, 2919–2931 (1964).
5. King, E. A. & Wagstaff, J. *Antarct. J. U.S.* **15**, 78–79 (1980).
6. MacKinnon, I. D. R. *et al. J. geophys. Res.* **87**, A413–A421 (1982).
7. Clanton, U. S. *et al. Cosmic Dust Catalog*, Vol. 3, 1 (Lyndon B. Johnson Space Center, Houston, 1982).
8. Wright, F. W. & Hodge, P. W. *Ann. N.Y. Acad. Sci.* **119**, 1, 1–361 (1964).
9. Wright, F. W. & Hodge, P. W. *J. geophys. Res.* **70**, 16, 3889–3898 (1965).
10. Heiken, G. & Wohletz, K. *Volcanic Ash*, University of California Press, Berkeley (1985).
11. Heiken, G. & Lofgren, G. *Bull. geol. Soc. Amer.* **82**, 1045–1050 (1971).
12. Baker, G. *Miner. Mag.* **36**, 1012–1023 (1967).
13. Sébastien, P., Gaudichet, A. *et al. J. microsc. spectrosc. Electron.* **5**, 83–97 (1980).
14. Gaudichet, A. *J. Phys. C2* **45**, Suppl. 2, 789–792 (1984).
15. Lefèvre, R. *et al. C.r. hebdom. Séanc. Acad. Sci., Paris* **301**, 1433–1438 (1985).
16. Cliff, G. & Lorimer, G. W. *J. Microsc.* **103**, 203–207 (1975).
17. Buat-Ménard, P. & Arnold, M. *Geophys. Res. Lett.* **5**, 4, 245–248 (1978).
18. Faivre-Pierret, R. *et al. Bull. volc.* **43**, 473–485 (1980).
19. Roy, W. R. *et al. Illinois State Geol. Surv. Environ. Geol. Not.* **96**, 1–69 (1981).
20. Ramsden, A. R. & Shibaoka, M. *Atmos. Environ.* **16**, 9, 2191–2206 (1982).
21. Reid, A. M. *et al. Eos* **54**, 606–607 (1973).
22. Heiken, G. H. *et al. Geochim. cosmochim. Acta* **38**, 1703–1718 (1977).

23. Heiken, G. H. & McKay, D. S. *Proc. 8th Lunar Sci. Conf.* 3243–3255 (1977).
24. Wohletz, K. M. & McQueen, R. G. *Geology* **12**, 591–594 (1984).
25. Wohletz, K. M. *J. Volcan. geotherm. Res.* **17**, 31–63 (1983).
26. Cristofolini, R. & Romano, R. *Mem. Soc. geol. It.* **23**, 99–115 (1982).
27. Clocchiatti, R. *et al. Bull. PIRPSEV, CNRS, Paris* **104**, 1–8 (1985).
28. Clocchiatti, R. & Metrich, N. *Bull. PIRPSEV, CNRS, Paris* **92**, 1–13 (1984).
29. Trichet, J. & Sella, C. *C.r. hebdom. Séanc. Acad. Sci., Paris* **267**, 1084–1086 (1968).
30. Kullerød, G. & Yoder, H. S. Jr *Yb Carnegie Instn Wash.* **63**, 218–222 (1964).
31. Deuser, W. G. *et al. Nature* **305**, 210–218 (1983).
32. Hulet, L. D. Jr *et al. Science* **210**, 1356–1358 (1980).
33. McGetchin, T. R. & Head, J. W. *Science* **180**, 68–70 (1973).
34. McGetchin, T. R. *et al. J. geophys. Res.* **79**, 3257–3272 (1974).

Sewage sludge as an acidity filter for groundwater-fed lakes

W. Davison

Freshwater Biological Association, The Ferry House, Ambleside LA22 0LP, UK

The acidification of lakes by acid rain has increased awareness of the problems associated with acid waters¹, and has stimulated work on self-neutralization mechanisms² and liming strategies³. However, some very acid lakes, particularly those which have been newly formed by the widespread extraction of coal, mineral ores and sand and gravel, owe their acidity to another source. If pyrite is present, as it often is, its exposure to air induces microbially mediated oxidation, resulting in the production of sulphuric acid⁴. The waters of the lakes then resemble dilute acid, typically with $\text{pH} < 3$, with an impoverished flora and fauna. Neutralizing such waters with lime ameliorates the situation, but only temporarily because further supply of acidic groundwater inexorably lowers the pH . Here I report the use of a combined treatment of lime and sewage sludge, which has produced a self-regulating system, capable of maintaining a stable pH . The organic material which is spread over the bottom of the lake acts as a chemical filter, removing acidic sulphate as it enters, and converting it to neutral sulphide.

The use of sewage sludge for lake reclamation was tested in a pilot scheme. British Industrial Sand Ltd provided a small ($3.6 \times 10^4 \text{ m}^2$) lake, Blue Lagoon, which had been formed by excavating sand below the water table at their Leziate site, near King's Lynn, Norfolk, England. It had no inflows, negligible catchment, a maximum depth of 4.5 m and a mean depth of 1.5 m. Over the past 10 years, during which time it had remained undisturbed, its waters had been consistently acid ($\text{pH} \sim 3$). Analysis showed that the major acid components were H_2SO_4 , 0.35 mM, $\text{Al}_2(\text{SO}_4)_3$, 0.11 mM and $\text{Fe}(\text{OH})\text{SO}_4$, 0.05 mM.

On 3 April 1984, 477 m^3 of activated sewage sludge containing 4.2% solids was pumped into the lake through a pipeline which was moved in an arc to cover approximately one quarter of the lake area. On 30 April 1984, 16.98 tonnes of calcium hydrated lime, equivalent to 15.3 tonnes of calcium hydroxide, were added through a submerged pipe to the deepest point.

Within days the sewage sludge settled out to leave clear water. The pH quickly rose to ~ 10 on adding the lime, after which, as carbon dioxide was supplied from the atmosphere, the solubility of calcium carbonate became the controlling factor and the pH stabilized between 7 and 8. To assess the pH without the effects of supply and removal of carbon dioxide by respiration and photosynthesis, it was also recorded after bubbling the sample with air for ≥ 30 min (Fig. 1). After two years the lake was still neutral.

Many factors may affect the pH , including the residence time of the water and further dissolution of precipitated calcium carbonate. The real test of the scheme is whether, first, sulphate is reduced and removed as sulphide, and, second, the rate of removal of sulphate is sufficient to cope with the further supply of acid from groundwater.

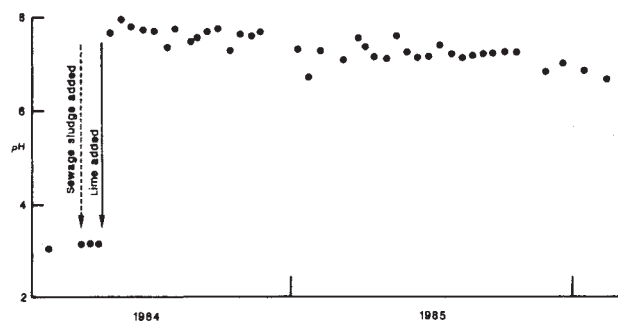


Fig. 1 Air-equilibrated pH of the surface waters of Blue Lagoon, plotted against time. The neutralization reactions involved are: $\text{H}_2\text{SO}_4 + \text{Ca}(\text{OH})_2 = \text{CaSO}_4 + 2\text{H}_2\text{O}$; $\text{Al}_2(\text{SO}_4)_3 + 3\text{Ca}(\text{OH})_2 = 2\text{Al}(\text{OH})_3 + 3\text{CaSO}_4$; $\text{Fe}(\text{OH})\text{SO}_4 + \text{Ca}(\text{OH})_2 = \text{Fe}(\text{OH})_3 + \text{CaSO}_4$.

On 27 August 1985, four cores of sediment and overlying water were collected from the deepest site using a Jenkin corer⁵. Analysis of two of the sediment cores revealed contents of organic carbon (11–20%) and acid volatile sulphide (1.2–18 mg per g dry weight) which were higher than would be normally expected for a very productive lake⁶. Evidently, appreciable quantities of sulphate had been reduced.

To quantify the rate and extent of sulphate reduction, two sealed sediment–water cores were incubated in the dark, initially at 17 °C, and then, after replenishing the sulphate-depleted water, at lower temperatures. Analysis of sulphate in the overlying water showed that sulphate was being removed at a rapid rate (Fig. 2), 835 $\text{mg m}^{-2} \text{ day}^{-1}$ at 17 °C and 278 $\text{mg m}^{-2} \text{ day}^{-1}$ at 10 and 4 °C. Knowing the original sulphate concentration of the lake water ($\sim 125 \text{ mg l}^{-1}$), and assuming it is the same in the groundwater, one can calculate the minimum water residence time required for the reduction of sulphate to sulphide in the sediment to be able to cope with additional supply of acid from the groundwater. At the rate of consumption for 17 °C the minimum residence time is 6 months; at the lower temperatures it is 1.5 yr. As the mean annual temperature will be within this range, the residence time must exceed 1 yr if the chemical filter is to keep pace with the input of acid. There is evidence to show that at the outset of our study the residence time of Blue Lagoon was ~ 1 yr, but that it has since progressively increased to several years, presumably due to organic matter clogging the pores of the sand. If reduction of sulphate occurs as efficiently in the lake as in the laboratory experiment, this mechanism alone will prevent re-acidification. When these incubation results are applied to the whole lake, they are equivalent to an annual base production of 0.7–2.1 mequiv. l^{-1} , which agrees with the figure of 1 mequiv. l^{-1} for hypolimnetic water².

So far only the consumption of freshly entering sulphate has been considered, but there is also sulphate already present within the water. Its associated acid was neutralized by the initial lime, so there is no requirement for this sulphate to be consumed to maintain neutrality. The concentration of sulphate, normalized with respect to sodium to compensate for evaporation and dilution effects, is highly variable, but did undoubtedly decline over a 2-yr period (Fig. 3). This indicates that removal of sulphate from the sediment was more than balancing further supply.

If the only supply of organic matter were from sewage sludge, the sediment's capacity for sulphate reduction would soon be exhausted. However, liming and the nutrients associated with the sludge stimulated high biological productivity, so that within a few months of treatment, algal populations were typical of a eutrophic pond. Thus, additional carbon, originating as atmospheric CO_2 , is continually supplied to the lake and eventually accumulates as organic-rich sediment. Photosynthesis is effectively used, through artificially induced eutrophication, to combat acidification.

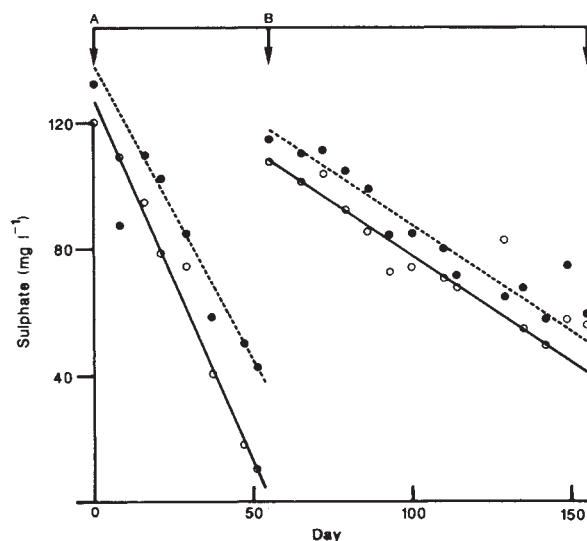


Fig. 2 Sulphate concentrations in the waters overlying the sediment of two sealed cores incubated in the dark. Initially (A–B) they were maintained at 17 °C, after which (at point B) the overlying waters were exchanged, under anoxic conditions, with water from the lake that had been deoxygenated using an Ar– CO_2 gas mixture to leave the pH unaltered. One core (O) was then incubated at 10 °C and the other (●) at 4 °C.

This scheme of permanent neutralization will be most effective when applied to large lakes with flat bottoms and deep water. Deep water (>5 m) prevents disturbance of the organic-rich sediment by wind-induced water movements, and flat bottoms help to ensure that the covering is uniform. Reduction of sulphate occurs readily in a sediment overlain by oxygenated water⁷, but it proceeds with greater efficiency if the overlying waters become anoxic, a situation encouraged by depths >5 m. Covering the sediment with reducing organic material cuts off an important supply of oxygen to the groundwater. This limitation of oxygen, which will decrease the rate of production of acid from pyrite, will be more effective, the greater the area of the lake.

Unfortunately, the only experimental site available, Blue Lagoon, is smaller and shallower than ideally required, according to the above criteria. It is not surprising, then, that the pH appears to be slowly declining (Fig. 1), indicating that the lake may return to an acidic state. However, sulphate has been consumed by the sewage sludge and photosynthesis will continue the process, so the new equilibrium pH which will event-

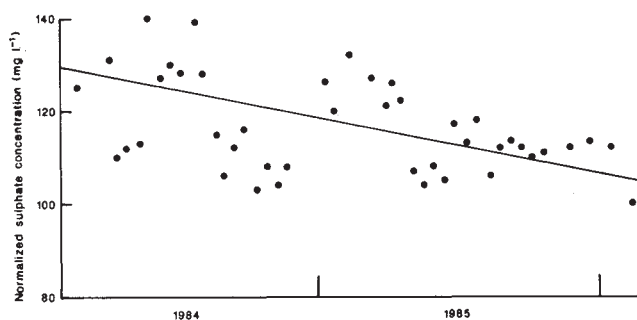


Fig. 3 Concentration of sulphate in Blue Lagoon, normalized with respect to the concentration of sodium. Normalized SO_4^{2-} concentration = $[\text{SO}_4^{2-}]_t \times [\text{Na}^+]_{\text{in}} / [\text{Na}^+]_t$, where the subscripts 'in' indicates initial concentration, and t the concentration at time t . Acidity removal is accomplished by the reduction of sulphate and ferric iron by organic material, and their precipitation as ferrous sulphide: $\text{SO}_4^{2-} + 2\text{H}^+ + \text{Fe}(\text{OH})_3(\text{s}) + 9/4 \text{CH}_2\text{O} = \text{FeS}(\text{s}) + 9/4 \text{CO}_2 + 19/4 \text{H}_2\text{O}$.

ually be established will be higher than the original value. Thus, the feasibility of using organic material as a chemical filter for acidity has been established. The effectiveness of the scheme will vary from lake to lake, depending on the relative fluxes of acid and base. For lakes of appropriate morphology this scheme offers the possibility of permanent neutralization from one initial treatment.

One of the main features of this scheme is the use of organic matter to provide a chemical filter to neutralize the supply of acid from groundwater. Lakes affected by atmospheric acidification are not usually supplied with acid groundwater, except perhaps those in a peat terrain. However, introducing organic carbon, by adding sewage sludge or by adding nutrients to stimulate primary productivity, may still effectively combat acidification. Even in acidic conditions⁸, the decomposition of organic material will introduce base². The level of base production measured in this and other studies², ~ 1 mequiv. $l^{-1} yr^{-1}$, is large compared with the acidity of rainwater, which at pH 4 is 0.1 mequiv. l^{-1} . Evidently, by artificially inducing eutrophication, sufficient base can be produced to neutralize acidification from atmospheric sources. Further work is urgently needed to test such schemes and to compare them with liming strategies.

I thank BIS Ltd for sponsoring this work and many BIS and FBA staff for their enthusiastic involvement.

Received 7 April; accepted 11 June 1986.

1. Mason, J. & Seip, H. M. *Ambio* 14, 45-51 (1985).
2. Cook, R. B., Kelly, C. A., Schindler, D. W. & Turner, M. A. *Limnol. Oceanogr.* 31, 134-148 (1986).
3. Fraser, J. E. & Brit, D. L. Preprint, General Research Corporation, McLean, Virginia (1982).
4. Stumm, W. & Morgan, J. J. *Aquatic Chemistry* (Wiley, New York, 1981).
5. Mortimer, C. H. *Limnol. Oceanogr.* 16, 387-404 (1971).
6. Davison, W., Lishman, J. P. & Hilton, J. *Geochim. cosmochim. Acta* 49, 1615-1620 (1985).
7. Kelly, C. A. & Rudd, J. W. M. *Biogeochemistry* 1, 63-77 (1984).
8. Herlihy, A. T. & Mills, A. L. *Appl. envir. Microbiol.* 49, 179-186 (1985).

Climatic influence on the isotopic composition of bone nitrogen

Tim H. E. Heaton, John C. Vogel,
Gertrud von la Chevallier & Gill Collett

Natural Isotopes Division, National Physical Research Laboratory,
CSIR, Pretoria, South Africa

The $^{13}C/^{12}C$ isotope ratios in animal¹ and human² bone can be used as indicators of diet, more recently it was shown that the $^{15}N/^{14}N$ ratios of animals and humans are similarly determined by the food they eat³⁻⁵. Specifically, the stable carbon isotope composition reflects the proportion of C_3 and C_4 plants at the base of the food chain^{1,2}, while both ^{15}N and ^{13}C reveal the difference between a marine and terrestrial diet in modern as well as archaeological contexts⁵⁻⁷. Here we present data for human and animal bones from southern Africa which only partly conform to previously recognized patterns for $^{15}N/^{14}N$ ratios. Prehistoric human bones from a particular coastal region of South Africa show $^{15}N/^{14}N$ ratios consistent with the marine and terrestrial diets indicated by the $^{13}C/^{12}C$ ratios, but bones of both prehistoric humans and modern wild animals from a larger part of the subcontinent show variations in $^{15}N/^{14}N$ ratios which cannot be ascribed to known variations in diet. It appears that, in some environments, nitrogen isotope studies must also take into account the possible influence of the climate.

Collagen or gelatine was isolated from the bones of protohistoric and prehistoric human skeletons and modern wild herbivorous animals from a variety of climatic/vegetational habitats in South Africa and Namibia (Fig. 1). Isotopic ratios were determined on N_2 and CO_2 produced by combustion of the

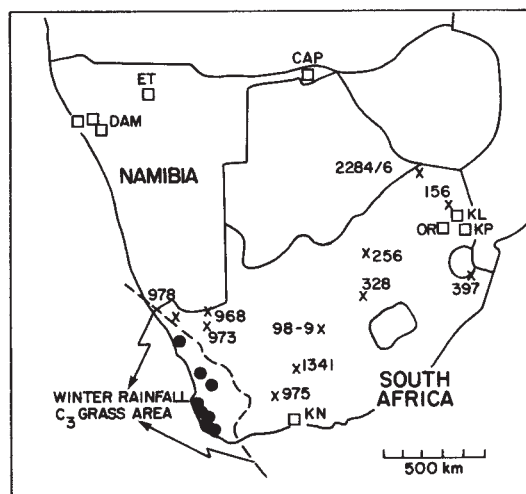


Fig. 1 Locations from which bone samples were obtained. ●, Human skeletons from sites 0-60 km from the south-west coast. x, Human skeletons from sites 60-500 km from the coast, laboratory sample numbers are shown. □, Herbivorous terrestrial animals from game reserves as indicated (CAP = Caprivi; DAM = Damaraland; ET = Etosha; KL = Klaserie; KN = Knysna; KP = Krugers Park; OR = Origstad dam).

isolates⁸, and on N_2 produced by the Kjeldahl digestion-lithium hypobromite method⁹. The ratios are reported in the usual $\delta^{15}N$ and $\delta^{13}C$ notation versus atmospheric N_2 and PDB carbonate standards, respectively. In all cases the N content of the samples ranged between 14% and 17%, within 10% of the expected value for protein. We therefore do not expect a noticeable postmortem shift in the isotopic values¹⁰.

The isotopic data indicated by the solid circles in Fig. 2a are for human bones from sites in the southwestern part of South Africa, an area with winter rainfall in which C_3 grasses (and other C_3 plants) predominate¹¹. Sealy and van der Merwe¹², presenting carbon isotope data for a set of samples from the same region, suggest that in this C_3 area a higher $\delta^{13}C$ value in the skeletons reflects a higher marine/terrestrial food ratio in the prehistoric human diet. Data for human bones from other parts of the world indicate that a larger proportion of marine food in the diet should also be reflected by higher $\delta^{15}N$ values⁶. The solid circles in Fig. 2a in fact show a close correlation of increasing $\delta^{15}N$ with increasing $\delta^{13}C$, and it appears that in this region the nitrogen isotope data for human skeletons independently confirm the marine/terrestrial dietary interpretation of the carbon isotopes.

Human skeletons from a large area in the interior of South Africa (crosses in Fig. 2a) have $\delta^{15}N$ values similar to those of skeletons from the south-west coast, extending into the range of $\delta^{15}N$ for a marine diet. The samples from the interior, however, were from sites 60 to 500 km from the nearest coast and, whilst some of the nomadic people may have spent part of their life at the coast, these distances imply that terrestrial rather than marine food would have formed the major component of their diet. The unexpected range of $\delta^{15}N$ values for bones of humans from the interior is moreover also displayed by bones of herbivores feeding exclusively on terrestrial plants. Thus bones of a single herbivorous species, the African Elephant, have $\delta^{15}N$ values from +2.6 to +14.6‰—a range which is wider than that previously reported for terrestrial herbivores, and which extends into the range of marine 'herbivores' (Fig. 2c).

Human bones from the interior and all the animal bones were from areas in which C_4 grasses grow among C_3 shrubs and trees. The range in $\delta^{13}C$ for the human (Fig. 2a) and animal bones¹

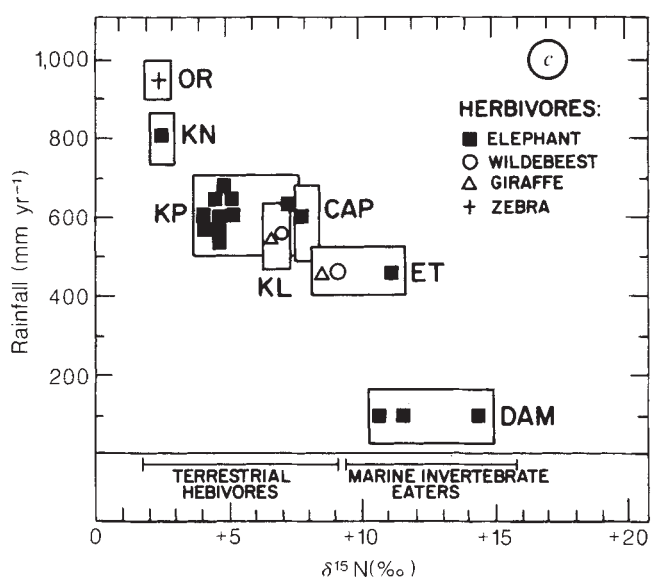
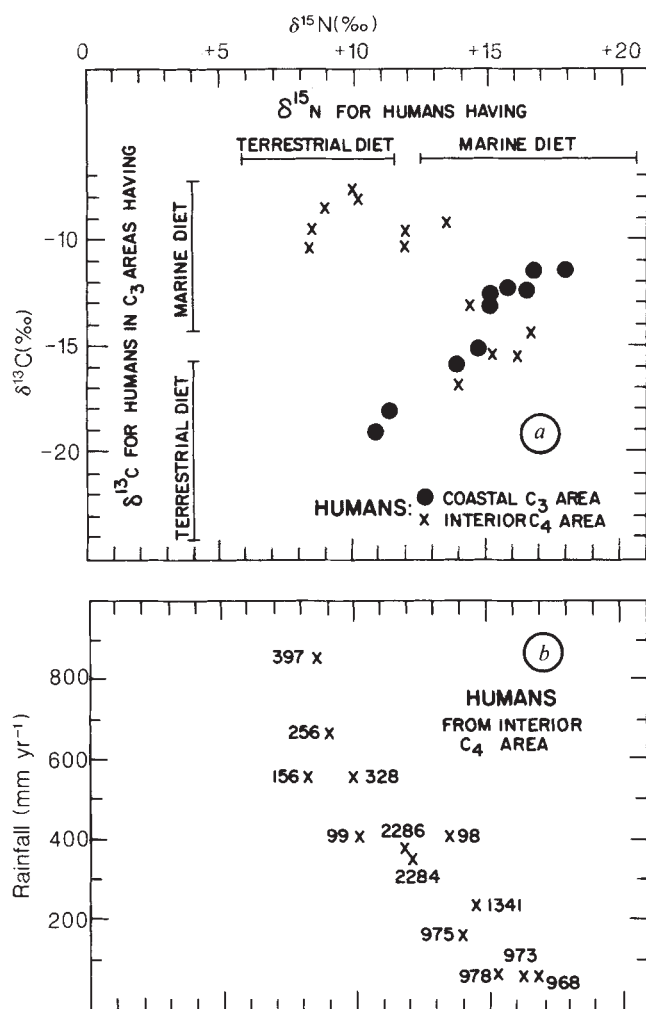


Fig. 2 Nitrogen isotope ratios of human and animal bones, $\delta^{15}\text{N}$ (in per mil) = $\{[(^{15}\text{N}/^{14}\text{N})_{\text{sample}}/(^{15}\text{N}/^{14}\text{N})_{\text{standard}}] - 1\} \times 10^3$; the standard is atmospheric N_2 . *a*, Nitrogen and carbon isotope ratios of humans (symbols as in Fig. 1) compared with the suggested range of isotopic compositions for humans having terrestrial and marine diets⁶. *b*, Nitrogen isotope ratios of humans from the interior vs. the mean annual rainfall of the collection site (numbers as in Fig. 1). *c*, Nitrogen isotope ratios of terrestrial herbivores versus the general range of mean annual rainfall for the collection area (letters refer to game reserves listed in Fig. 1). The horizontal bars show the published range of isotopic compositions for bones of terrestrial herbivores and marine invertebrate-eaters for comparison⁵.

in these areas reflects the different $^{13}\text{C}/^{12}\text{C}$ ratios for C₄ and C₃ plants and the variation in grazing versus browsing in the diet. There is no evidence, however, for a related effect on the nitrogen isotopes: C₄ and C₃ plants have similar $\delta^{15}\text{N}$ values¹³, and grazers (wildebeest) and browsers (giraffe) from the same area, whilst having very different $\delta^{13}\text{C}$ values¹, have very similar $\delta^{15}\text{N}$ values (Fig. 2c). Published analyses for the $\delta^{15}\text{N}$ values of modern nitrogen-fixing and non-fixing plants differ, on average, by only 2 to 3‰¹³; it therefore appears that the range in bone $\delta^{15}\text{N}$ values also cannot be explained in terms of different proportions of leguminous and non-leguminous plants in the diet.

We suggest that the wide range in $\delta^{15}\text{N}$ for human and animal bones in southern Africa reflects the very wide range of climatic habitats. In Fig. 2b and 2c the $\delta^{15}\text{N}$ values of the humans from the interior and of the herbivores are plotted against the mean annual rainfall for the sites from which the bones were collected. Bearing in mind the fact that some of the humans and animals are migratory (elephants moving between Etosha and Damara-land, for example), it nevertheless appears that there is a definite tendency for increasing ^{15}N with increasing aridity. Thus the animal or human bones with $\delta^{15}\text{N}$ values higher than +10‰ or +13‰ respectively, falling into the range of a 'marine diet' isotopic signature, are from areas receiving less than about 400 mm of rain per year (Fig. 2b, c).

The increase in nitrogen isotope fractionation that is observed in the bones of humans and particular species of ungulates living in arid environments must take place somewhere along the food chain. It occurs either in the plants growing under hot xeric

conditions or in the animals themselves. Measurements on C₃ and C₄ plants from a variety of habitats in southern Africa thus far show no obvious relationship of $\delta^{15}\text{N}$ with rainfall or photo-synthetic pathway. The higher $\delta^{15}\text{N}$ values of animals in arid areas are therefore more likely to be linked with the nitrogen metabolism in the body of the animal itself.

The data presented here indicate that the interpretation of the nitrogen isotope ratios of bones in terms of diet may be more complicated than previous studies suggest, and that the influence of climate also needs to be taken into account. $^{15}\text{N}/^{14}\text{N}$ ratios may in fact prove to be a useful tool for studying past climatic variations, especially in areas where large changes in precipitation have occurred.

We thank S. Braine, Möwe Bay, R. Thwaites, Saasveld and A. Hall-Martin, Skukuza, for supplying some of the elephant bone samples, and Siep Talma for supervising mass spectrometric analyses in the laboratory.

Received 31 December 1985; accepted 22 May 1986.

1. Vogel, J. C. S. *Afr. J. Sci.* **74**, 298-301 (1978).
2. Vogel, J. C. & van der Merwe, N. J. *Am. Antiq.* **42**, 238-242 (1977).
3. DeNiro, M. J. & Epstein, S. *Geochim. cosmochim. Acta* **45**, 341-351 (1981).
4. Minagawa, M. & Wada, E. *Geochim. cosmochim. Acta* **48**, 1135-1140 (1984).
5. Schoeninger, M. J. & DeNiro, M. J. *Geochim. cosmochim. Acta* **48**, 625-639 (1984).
6. Schoeninger, M. J., DeNiro, M. J. & Tauber, H. *Science* **220**, 1381-1383 (1983).
7. Tauber, H. *Nature* **292**, 332-333 (1981).
8. Schiegl, W. E. & Vogel, J. C. *Earth planet. Sci. Lett.* **7**, 307-313 (1970).
9. Heaton, T. H. E. & Collett, G. M. *CSIR Res. Rep.* 624 (1985).
10. DeNiro, M. J. *Nature* **317**, 806-809 (1985).
11. Vogel, J. C., Fuls, A. & Ellis, R. P. S. *Afr. J. Sci.* **74**, 209-215 (1978).
12. Sealy, J. C. & van der Merwe, N. J. *Nature* **315**, 138-140 (1985).
13. DeNiro, M. J. & Hastorf, C. A. *Geochim. cosmochim. Acta* **49**, 97-115 (1985).

Functions of the ON and OFF channels of the visual system

Peter H. Schiller, Julie H. Sandell
& John H. R. Maunsell

Whitaker College, Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

In the mammalian eye, the ON-centre and OFF-centre retinal ganglion cells form two major pathways projecting to central visual structures from the retina. These two pathways originate at the bipolar cell level: one class of bipolar cells becomes hyperpolarized in response to light, as do all photoreceptor cells, and the other class becomes depolarized on exposure to light, thereby inverting the receptor signal. It has recently become possible to examine the functional role of the ON-pathway in vision by selectively blocking it at the bipolar cell level using the glutamate neurotransmitter analogue 2-amino-4-phosphonobutyrate (APB)¹. APB application to monkey, cat and rabbit retinas abolishes ON responses in retinal ganglion cells, the lateral geniculate nucleus and the visual cortex but has no effect on the centre-surround antagonism of OFF cells or the orientation and direction selectivities in the cortex²⁻⁵. These and related findings⁶⁻¹¹ suggest that the ON and OFF pathways remain largely separate through the lateral geniculate nucleus and that in the cortex, contrary to some hypotheses, they are not directly involved in mechanisms giving rise to orientation and direction selectivities. We have examined the roles of the ON and OFF channels in vision in rhesus monkeys trained to do visual detection and discrimination tasks. We report here that the ON channel is reversibly blocked by injection of APB into the vitreous. Detection of light increment but not of light decrement is severely impaired, and there is a pronounced loss in contrast sensitivity. The perception of shape, colour, flicker, movement and stereo images is only mildly impaired, but longer times are required for their discrimination. Our results suggest that two reasons that the mammalian visual system has both ON and OFF channels is to yield equal sensitivity and rapid information transfer for both incremental and decremental light stimuli and to facilitate high contrast sensitivity.

Six monkeys were anaesthetized with halothane and given more than 50 vitreal injections by sterile procedures. The volume of APB used was 75 μ l (4–20 mM APB in sterile saline) to yield estimated concentrations of 100–500 μ M in the eye. The effects of such injections were assessed in separate experiments with the electroretinogram, since our previous studies had shown that when the ON channel is blocked as determined by single-cell recordings in the lateral geniculate nucleus, the b-wave of the electroretinogram is eliminated¹². Although there is some variability in the exact time-course of APB action with the intravitreal injection method, we found that a single injection was usually effective within 1 h and lasted 3–8 h. With the concentrations used in this study, recovery was complete by the following day. In all cases only one eye was injected and the other was patched.

Monkeys were deprived of water and rewarded with apple juice. Their eye movements were monitored with the scleral search coil technique, and they were rewarded for correctly making saccades to visual stimuli presented on a colour monitor. A PDP 11/73 computer was used to drive the monitor, control the experiment and record and store the eye movement and performance data. All trials began with the appearance of a small central spot. In the detection paradigm, after the animal had made saccades to this spot and maintained fixation on it for 400–800 ms, a target stimulus appeared at one of several randomly chosen locations, and the animal's task was to make saccades to it; the stimulus remained on until it was foveated.

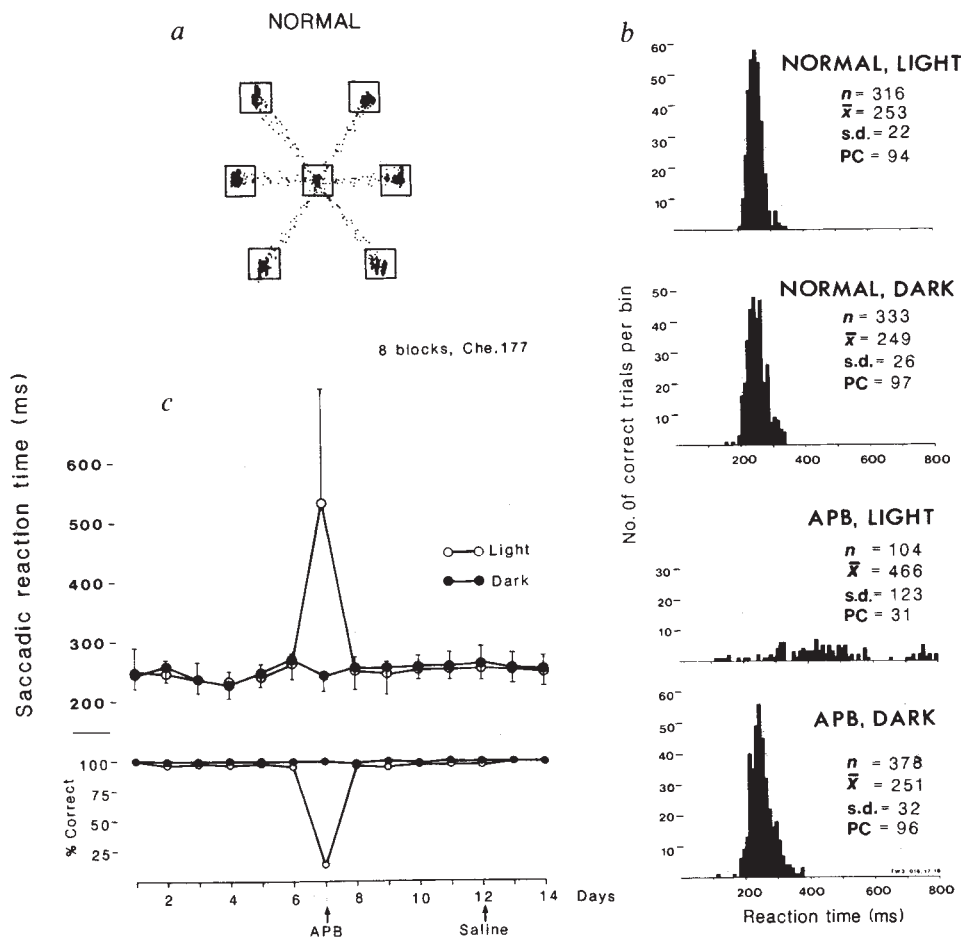
Trials on which saccades were made to locations other than the target were aborted. When the saccade brought the fovea to within 2° of the target centre, it was judged by the programme to be successful and the animal was rewarded with a drop of apple juice. This paradigm, which permits accurate placement of stimuli on selected portions of the retina, is learned rapidly and performed accurately. Trials can be run in close succession without fatigue (with inter-trial intervals of 2–3 s), and monkeys regularly perform 1,500–2,000 trials per day. This permits relatively speedy generation of psychophysical functions. Figure 1a shows eye movement data collected for 48 successive errorless trials in which a target appeared 8 times at each of 6 target locations, showing the normal animal's characteristic, disciplined performance.

To test the hypothesis that the ON channel contributes predominantly to the detection of stimuli seen by virtue of light increment while the OFF channel contributes to the detection of light decrement, we trained monkeys to make saccades to targets which were either lighter than or darker than the background. Figure 1b shows the distribution of saccadic latencies on a set of light and dark stimuli in the normal and APB-injected monkey. One to six hours after APB administration (75 μ l of 8 mM APB), there was a pronounced deficit in the detection of light incremental stimuli, accompanied by a large increase and spread in saccadic latencies. Performance on light decremental stimuli was unaffected. Figure 1c shows the monkey's performance over 14 successive days on this task, plotting the saccadic reaction times and the per cent correct performance on one set of light and dark stimuli. The APB was injected on day 7 during this sequence, showing the selective impairment for stimuli brighter than background, and also showing that saline injection of the same volume on day 12 did not affect the animal's performance. Similar effects were obtained using a range of target stimuli measuring 17–220 cd m^{-2} . We also obtained thresholds for the detection of light stimuli in dark-adapted animals: when the ON channel was blocked, detection thresholds were raised by 1.5–2.0 log units, suggesting that APB has a dramatic effect on the eye's rod system.

We examined several other perceptual functions for which we used discrimination rather than detection tasks. The paradigm was quite similar to that used for the detection of light increment and decrement, but instead of one target, several stimuli appeared after the central spot was fixated. One of these stimuli was different from the rest and the animal's task was to make saccades to this target. The physical characteristics of the stimuli could be varied systematically and were presented in randomized sequences. These tests assessed thresholds for the detection of colour stimuli of different degrees of saturation, the maximal temporal rate of flicker discrimination, minimal stereoscopic depth perception using random-dot stereograms, grating acuity, movement perception using random dot displays and the discrimination of gratings having different orientations. Only mild deficits were found on these tasks when the ON channel was blocked with APB. On all of the tasks, however, there was a consistent increase in the latency to performance, averaging 50 ms.

We also examined the contrast sensitivity of animals following APB administration and found considerable impairment using a variety of contrast sensitivity tests. Figure 2 shows the results from one of these. Monkeys were shown six stimuli, five of which were checkerboards while one was a homogeneous stimulus of the same flux. The animal was trained to make saccades to this homogeneous stimulus. Checkerboards of six different spatial frequencies were used, all at the same contrast, randomized over successive trials within a single block. Different contrasts were used in successive blocks. The stimuli appeared at equally spaced locations around the fixation spot at a distance of 8°. The results show a significant loss in contrast sensitivity, particularly at optimal spatial frequencies, as seen in both the per cent correct and the reaction time data. In other tests a

Fig. 1 *a*, An *x-y* frame showing traces of saccades made from a central fixation point to a target presented randomly at one of six positions. Targets are 8° eccentric and are squares subtending 1° of visual angle; data from 48 successive trials are shown. Boxes represent spatial limits set for correct acquisition. *b*, Distribution of saccadic latencies and per cent correct performance before and after APB injection ($75 \mu\text{l}$ of 8 mM) to light incremental and light decremental stimuli. Base illumination set at 121 cd m^{-2} . The light stimuli measured 186 and the dark 59 cd m^{-2} . *n*, Number of trials per histogram; $\bar{x}$, mean; *s.d.*, standard deviation. PC, per cent correct. *c*, Saccadic reaction times and per cent correct performance on 14 successive days for incremental (circles) and decremental (disks) stimuli. APB ($75 \mu\text{l}$ of 8 mM) was injected on day 7 and saline ($75 \mu\text{l}$) on day 12. Vertical bars show standard deviations, upward for incremental and downward for decremental stimuli.



two-alternative forced-choice staircase procedure was used in which sinusoidal gratings had to be discriminated from homogeneous stimuli; injections of $450 \mu\text{M}$ APB resulted in a loss of foveal contrast sensitivity of up to 25%.

These results suggest two possible reasons for the existence of both an ON and an OFF channel in the mammalian visual system. One is to provide a means for transmitting information about both light increment and light decrement with an excitatory process to the central nervous system. Since the maintained activity of retinal ganglion cells is relatively low, the information conveyed by a decrease in activity (for example, the effect of

light increment on OFF cells) is difficult to utilize effectively by central visual system neurones, particularly since their maintained activity is even lower than that of retinal ganglion cells. Having both an ON and an OFF channel therefore makes possible efficient information transfer for either sign of contrast change. This is necessary because of the high premium that exists on the speed of information processing in the animal kingdom, both for appetitive and for avoidance behaviours. The second reason for the existence of the ON and OFF channels is to provide increased contrast sensitivity, probably as a result of some form of push-pull interaction between the ON and OFF channels at higher levels in the visual system. High contrast sensitivity is undoubtedly a most desirable attribute for optimal visual function. Neither of these considerations identifies a unique solution to a biological problem. Rapid information transfer and high contrast sensitivity could probably also be accomplished by having a single channel with a rather high maintained discharge throughout the visual system. Such a solution, however, would probably require an unacceptably high rate of metabolic activity.

This research was supported by NIH grant EY00676, NSF grant BNS-8310399 and NRSA grant F32 NS06971-02. We thank David Wilson for his help in training the animals.

Received 28 February; accepted 20 May 1986.

1. Slaughter, M. M. & Miller, R. F. *Science* **211**, 182-184 (1981).
2. Schiller, P. H. *Nature* **297**, 580-583 (1982).
3. Horton, J. C. & Sherk, H. *J. Neurosci.* **4**, 374-380 (1984).
4. Sherk, H. & Horton, J. C. *J. Neurosci.* **4**, 381-393 (1984).
5. Knapp, A. G. & Mistle, L. A. *J. Neurophysiol.* **50**, 1236-1245 (1983).
6. Conway, J. L. & Schiller, P. H. *J. Neurophysiol.* **50**, 1330-1342 (1983).
7. LeVay, S. & McConnell, S. K. *Nature* **300**, 350-351 (1982).
8. Norton, T. T., Rager, G. & Kretz, R. *Brain Res.* **327**, 319-323 (1985).
9. Raczowski, D. & Fitzpatrick, D. *Neurosci. Abstr.* **68**(9), 227 (1985).
10. Stryker, M. P. & Zahs, K. R. *J. Neurosci.* **3**, 1943-1951 (1983).
11. Waessle, H., Boycott, B. B. & Illing, R. B. *Proc. R. Soc. B212*, 177-195 (1981).
12. Knapp, A. G. & Schiller, P. H. *Vision Res.* **24**, 1841-1846 (1984).

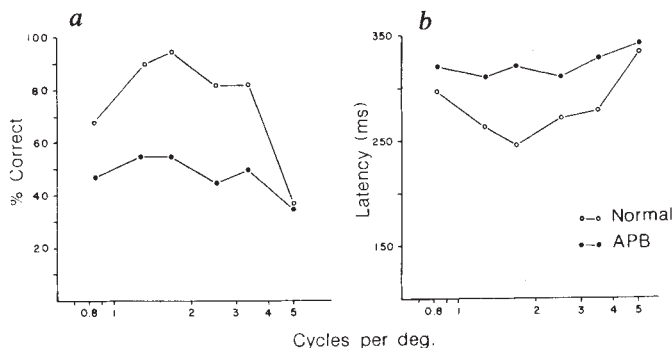


Fig. 2 Per cent correct (*a*) and mean saccadic latencies (*b*) on a discrimination task requiring an animal to make saccades to a homogeneous stimulus paired with a set of five checkerboard stimuli; sets were of different spatial frequencies but of the same contrast. Background illumination was 25 cd m^{-2} , the homogeneous stimulus was 68 cd m^{-2} and the high- and low-contrast checkerboards were 88 and 48 cd m^{-2} , respectively. Each point is based on 60 trials.

Expression of functional sodium channels from cloned cDNA

Masaharu Noda, Takayuki Ikeda, Harukazu Suzuki,
Hiroshi Takeshima, Tomoyuki Takahashi*,
Motoy Kuno* & Shosaku Numa

Departments of Medical Chemistry and Molecular Genetics
and * Physiology, Kyoto University Faculty of Medicine,
Kyoto 606, Japan

The voltage-gated sodium channel is a transmembrane protein essential for the generation of action potentials in excitable cells¹. It has been reported that sodium channels purified from the electric organ of the electric eel, *Electrophorus electricus*^{2,3}, and from chick cardiac muscle⁴ consist of a single polypeptide of relative molecular mass (M_r) ~260,000, whereas those purified from rat brain⁵ and from rat^{6,7} and rabbit skeletal muscle⁷ contain, in addition to the large polypeptide, one or two smaller polypeptides of M_r 33,000–43,000. The primary structures of the *Electrophorus* sodium channel⁸ and two distinct sodium channel large polypeptides⁹ (designated as sodium channels I and II) from rat brain have been elucidated by cloning and sequencing the complementary DNAs. The purified sodium channel preparations from *Electrophorus* electroplax¹⁰ and from mammalian muscle^{11,12} and brain^{13–15}, when reconstituted into lipid vesicles or planar lipid bilayers, exhibit some functional activities. The successful reconstitution with the *Electrophorus* preparation would imply that the large polypeptide alone is sufficient to form functional sodium channels. However, studies with the rat brain preparation suggest that the smaller polypeptide of M_r 36,000 is also required for the integrity of the saxitoxin (STX) or tetrodotoxin (TTX) binding site of the sodium channel¹⁶. Here we report that the messenger RNAs generated by transcription of the cloned cDNAs encoding the rat brain sodium channel large polypeptides, when injected into *Xenopus* oocytes, can direct the formation of functional sodium channels.

Recombinant plasmids that carry the bacteriophage SP6 promoter^{17,18}, linked to the entire protein-coding region of the cDNA for rat sodium channel I (pRI-4 with the 33-nucleotide insertion⁹ or pRI-3 without the insertion) or sodium channel II (pRII-2), were constructed (Fig. 2). Plasmid transcription *in vitro* by SP6 polymerase generated mRNAs specific for the respective sodium channel large polypeptides (Fig. 1). The estimated sizes of the mRNAs agree with those expected from the structure of the plasmids, except that smaller RNA species were contained in the preparations derived from pRI-3 and pRI-4.

Xenopus oocytes that had been injected with the sodium channel II-specific mRNA derived from pRII-2 and then incubated for 3–6 days showed a transient inward current when the holding membrane potential was shifted from –100 mV to –10 mV under voltage clamp. Following application of TTX (0.3 μ M), the inward current was progressively diminished and disappeared within a few minutes (Fig. 3a). Similarly, STX (0.3 μ M) blocked the inward current completely. The effects of both the toxins were reversible. In Ringer's solution, the maximum inward currents induced by the sodium channel II-specific mRNA ranged up to 26 μ A, being much larger than those observed in oocytes injected with poly(A)⁺ RNA from rat brain (up to 1.5 μ A) (see also refs 19, 20). On the other hand, oocytes injected with the sodium channel I-specific mRNA showed only a small response; the mRNA derived from pRI-3 evoked maximum inward currents up to ~50 nA, which were TTX-sensitive, whereas the mRNA derived from pRI-4 induced no detectable currents. Because of the low signal-to-noise ratio for the response evoked by the sodium channel I-specific mRNAs, we restricted our further analysis to the current induced

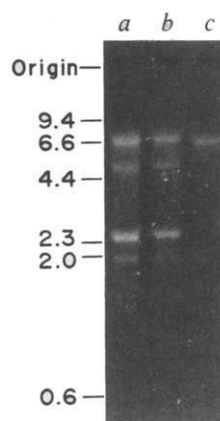


Fig. 1 Electrophoretic analysis of the mRNAs synthesized *in vitro* using pRI-3 (a), pRI-4 (b) or pRII-2 (c) as template. Samples of the mRNAs (5 μ g in a, b; 1.3 μ g in c) were denatured with 1 M glyoxal and 50% dimethyl sulphoxide³³ and electrophoresed on a 1.5% agarose gel in 10 mM sodium phosphate buffer (pH 7.0). RNA was visualized by staining with ethidium bromide. The size markers used were the *Hind*III cleavage products of phage λ DNA (sizes in kilobases). The estimated size of the largest mRNA species on each lane agrees with the expected size (~6.7, ~6.7 and ~6.5 kilobases, respectively, in a, b and c).

by the sodium channel II-specific mRNA and compared its properties with those of the current induced by poly(A)⁺ RNA from rat brain.

Figure 3b shows an example of the dose-response curves for TTX obtained from oocytes injected with the sodium channel II-specific mRNA. The apparent dissociation constant for TTX (K_{TTX}) ranged from 10 nM to 14 nM. This was comparable to the K_{TTX} value (6–7 nM) obtained for the sodium channels produced in oocytes injected with rat brain poly(A)⁺ RNA; the latter value is similar to that reported previously (9–10 nM)^{19,20}.

When the external Na⁺ concentration was lowered by replacement with tetraethylammonium, tetramethylammonium or sucrose, the TTX-sensitive inward current was reduced in a dose-dependent manner, being virtually abolished at ~3 mM Na⁺ (Fig. 4a). The fast Na⁺ current observed in the injected oocytes differed clearly from the persistent Na⁺ current which can be induced in native oocytes only by prolonged depolarization and is blocked by TTX only at a dose of ~0.5 mM (refs 21, 22). The fast Na⁺ current was not detected in non-injected oocytes. Figure 4b exemplifies current recordings in Ringer's solution that were made when the oocyte membrane potential was stepped from –100 mV to various levels between –51 mV and +59 mV. In these records, the TTX-insensitive components were subtracted. The peak current-voltage (*I*–*V*) relationship obtained by this procedure is shown in Fig. 4c. The maximum current occurred in a range of –12 to +2 mV. This range was similar to that observed for the Na⁺ current measured in oocytes injected with rat brain poly(A)⁺ RNA (–8 to 0 mV) (see also refs 19, 20). From the internal Na⁺ concentration measured in *Xenopus* oocytes^{23,24}, the reversal potential of the Na⁺ current would be expected to be 40–66 mV. However, the inward current was not reversed at potential levels up to 80 mV. A decrease in the external Na⁺ concentration from 118 mM to 80 mM reduced the inward current, but again failed to reverse the current. A significant permeation of Ca²⁺ through the TTX-sensitive channel was unlikely because the inward current was not affected by replacement of external Ca²⁺ with equimolar Mn²⁺ or by an increased Ca²⁺ concentration (20 mM). The TTX-sensitive inward current observed in oocytes injected with rat brain poly(A)⁺ RNA was also unaffected by Ca²⁺-free saline, in

Fig. 2 Construction of recombinant plasmids used for expression of rat sodium channel I (pRI-3) (a) and sodium channel II (pRII-2) (b). Only relevant restriction endonuclease sites are shown and identified by numbers indicating the 5'-terminal nucleotide generated by cleavage (for nucleotide numbers, see ref. 9); endonuclease sites that no longer exist in pRI-3 or pRII-2 and those corresponding to the ends of the synthetic linkers (indicated by narrow open boxes) are parenthesized. Protein-coding regions are represented by solid boxes, 5'- and 3'-noncoding regions and vector sequences immediately following the poly(dA)·poly(dT) tract by solid lines, and poly(dA)·poly(dT) tracts by wavy lines. The cDNA clones⁹ used for the construction are indicated by thick lines, and their fragments constituting pRI-3 or pRII-2 by bars underneath. The 33-nucleotide insertion in prSCH71 is shown by an inverted triangle. A scale of 1 kilobase pair (kb) is given.

Methods. The pSP64 (ref. 18) recombinants carrying the cDNA for rat brain sodium channel I with (pRI-4) or without the 33-nucleotide insertion (pRI-3) were constructed as follows. To delete the ATG triplet composed of nucleotides -8 to -6 in the 5'-noncoding region, the synthetic oligonucleotide linker* was prepared using an automatic DNA synthesizer (Applied Biosystems). The ~1.9-kb *FokI*(5)/*FokI* (on vector DNA) fragment from prSCH74 was ligated with the synthetic linker and cleaved with *EcoRI*. The ~1.6-kb fragment formed was isolated and cleaved with *PstI*. The resulting *PstI*/*EcoRI* fragment and the ~2.7-kb *EcoRI*/*PstI* fragment from pUC18 (ref. 30) were ligated to yield pURI-1. The *PstI*(-74)/*EcoRI*(1,556) fragment from prSCH74, the *EcoRI*(1,556)/*AvaII*(1,967) fragment from prSCH71, the *AvaII*(1,967)/*BamHI*(2,116) and *BamHI*(2,116)/*HpaII*(3,264) fragments from prSCH28, the *HpaII*(3,264)/*HindIII*(3,534) fragment from prSCH109 and the ~3.0-kb *HindIII*/*PstI* fragment from pSP65 (ref. 18) were ligated to yield pSRI-A, which was used to construct pRI-3; for the construction of pRI-4, the *EcoRI*(1,556)/*BamHI*(2,116) fragment from prSCH71, instead of the *EcoRI*(1,556)/*AvaII*(1,967) fragment from prSCH71 and the *AvaII*(1,967)/*BamHI*(2,116) fragment from prSCH28, was used to yield pSRI-B. The *PvuII*(5,498)/*PvuII*(6,251) fragment from prSCH109 was inserted into the *SmaI* site of pUC12 (ref. 31) in the same orientation as the *lacZ* promoter to yield pURI-2. The *EcoRI*/*HindIII* fragment of 1,945 (or 1,978) base pairs (bp) from pSRI-A (or pSRI-B), the *HindIII* (3,534)/*AvaII*(5,557) fragment from prSCH109, the 703-bp *AvaII*/*XbaI* (on vector DNA) fragment from pURI-2 and the ~3.0-kb *XbaI*/*EcoRI* fragment from pUC12 were ligated to yield pURI-A (or pURI-B). The 1,565-bp *PstI*/*EcoRI* fragment from pURI-1, the 4,671-bp (or 4,704-bp) *EcoRI*/*XbaI* fragment from pURI-A (or pURI-B), the ~0.4-kb *XbaI*(7,887)/*PvuII* (on vector DNA) fragment from prSCH109 and the ~3.0-kb *HincII*/*PstI* fragment from pSP64 were ligated to produce pRI-3 (or pRI-4). The pSP65 recombinant carrying the cDNA for rat brain sodium channel II (pRII-2) was constructed as follows. The synthetic oligonucleotide linker† was prepared to delete the ATG triplet composed of nucleotides -8 to -6 in the 5'-noncoding region. The *FokI*(5)/*FokI*(1,040) fragment from prSCH93 was ligated with the synthetic linker and cleaved with *PstI*. The ~0.8-kb fragment formed was isolated and cleaved with *EcoRI*. The resulting *EcoRI*/*PstI* fragment and the ~2.7-kb *PstI*/*EcoRI* fragment from pUC19 (ref. 30) were ligated to yield pURI-1. The *EcoRI*(-134)/*NcoI*(1,275) and *NcoI*(1,275)/*PstI*(1,396) fragments from prSCH93, the *PstI*(1,396)/*HindIII*(3,102) fragment from prSCH78, the *HindIII*(3,102)/*SacI*(3,931) fragment from prSCH82 and the ~3.0-kb *SacI*/*EcoRI* fragment from pSP64 were ligated to yield pSRII-1. The *HindIII*(3,102)/*SacI*(3,931) fragment from prSCH82, the *SacI*(3,931)/*EcoRI*(5,869) and *EcoRI*(5,869)/*Sau3A1*(6,084) and ~0.3-kb *BglII*(7,554)/*PvuII* (on vector DNA) fragments from prSCH202 and the ~3.0-kb *HincII*/*HindIII* fragment from pSP65 were ligated to yield pSRII-2. The 733-bp *EcoRI*/*ApaI* fragment from pURI-1, the 2,409-bp *ApaI*/*BglII* fragment from pSRII-1, the ~3.2-kb *BglII*/*XbaI* (on vector DNA) fragment from pSRII-2 and the ~3.0-kb *XbaI*/*EcoRI* fragment from pSP65 were ligated to produce pRII-2. The mRNAs specific for rat sodium channels I and II were synthesized *in vitro*^{17,18}, using *SmaI*-cleaved pRI-3 (or pRI-4) and *SacI*-cleaved pRII-2, respectively, as templates. Transcription primed³² with the cap dinucleotide m⁷G(5')ppp(5')G (1 mM).

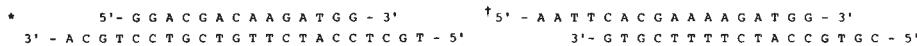
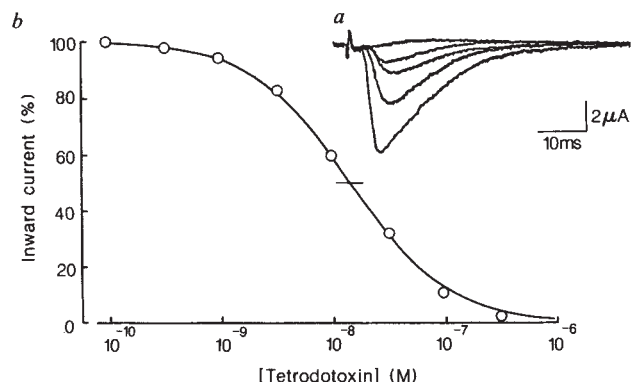


Fig. 3 Effect of TTX on depolarization-activated whole-cell inward currents in *Xenopus* oocytes injected with the sodium channel II-specific mRNA. **a**, Time course. The inward currents were elicited by a 50-ms pulse stepped from -100 mV to -10 mV. The five records (from bottom to top) were obtained before and 15 s, 20 s, 35 s and 70 s after superfusion with 0.3 μ M TTX, respectively. The response obtained 10 min after washing the TTX solution was 88% of that observed before TTX application. Transient capacitive currents were roughly compensated by a network with three adjustable time constants. **b**, Dose-response curve obtained from another oocyte. The peak inward current evoked by a step from -100 mV to -10 mV, relative to that observed before TTX application, is plotted against the logarithm of TTX concentration. Each point was obtained 3 min after exposure to a new dose of TTX. The measurements were made in the direction of increasing TTX concentration. The record obtained with ~1 μ M TTX was subtracted from the control and those obtained with other doses of TTX. The curve represents the dose-response relation expected from a K_{TTX} of 14 nM (indicated by a horizontal bar) according to the equation $y = (1 + T/K_{\text{TTX}})^{-1}$, where T is the TTX concentration.

Methods. Total RNA was extracted from the whole brain of male Wistar rats (~200 g body weight) by the phenol method³⁴, and poly(A)⁺ RNA was isolated by oligo(dT)-cellulose chromatography³⁵. Fractionation of poly(A)⁺ RNA by sucrose gradient centrifugation³⁶ was performed as follows. Poly(A)⁺ RNA (100 μ g per tube) was centrifuged in linear sucrose gradients (15–30%, w/v) in a Beckman SW-41 rotor at 40,000 r.p.m. for 16 h at 4°C and 25 fractions were collected from each gradient. The fractions sedimenting more slowly than 28S ribosomal RNA and the remaining fractions were each pooled, extracted with phenol/chloroform, precipitated with ethanol and dissolved in water. *Xenopus laevis* oocytes were injected with the following RNA preparations separately or in combination: sodium channel I-specific mRNA with or without the 33-nucleotide insertion (1.0 μ g μ l⁻¹), the sodium channel II-specific mRNA (0.2 μ g μ l⁻¹), poly(A)⁺ RNA (1.0 μ g μ l⁻¹), the slow-sedimenting fraction (<28 S) of poly(A)⁺ RNA (2.4 μ g μ l⁻¹) and the fast-sedimenting fraction ($\geq$ 28 S) of poly(A)⁺ RNA (1.6 μ g μ l⁻¹). The average volume injected was ~50 nl per oocyte. The injected oocytes were incubated at 19°C for 3–6 days in modified Barth's medium³⁷ containing gentamicin (0.1 mg ml⁻¹) and mycostatin (15 μ g ml⁻¹). The follicular cell layer was removed^{26,38} from oocytes prior to electrophysiological measurements. Whole-cell currents were recorded with a two-microelectrode voltage clamp at 20 $\pm$ 2°C in Ringer's solution of the following composition (in mM) unless otherwise specified: NaCl, 115; KCl, 2.5; CaCl₂, 1.8; HEPES (pH 7.4), 5 (containing 2.75 mM Na⁺). The voltage recording electrode was filled with 3 M KCl. The current electrode was filled with a solution of (in mM): CsF, 250; CsCl, 250; EGTA, 50; HEPES (pH 7.4), 10.



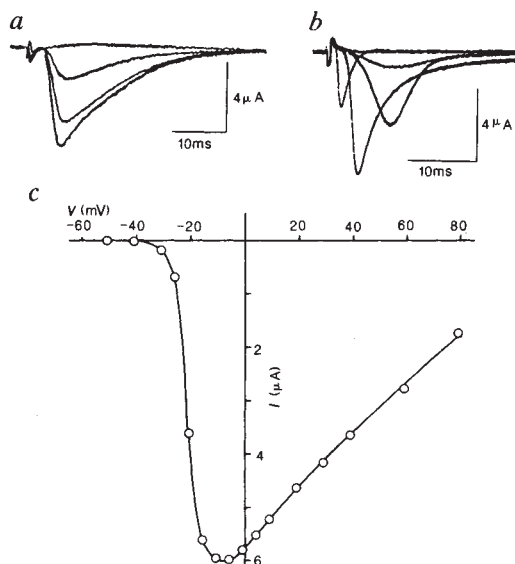


Fig. 4 Properties of depolarization-activated whole-cell inward currents in oocytes injected with the sodium channel II-specific mRNA. **a**, Effects of changes in external Na^+ concentration replaced by tetraethylammonium ions. The inward currents were elicited by a 50-ms pulse stepped from -100 mV to -10 mV. The four records (from bottom to top) were obtained at external Na^+ concentrations of 118 mM, 80 mM, 41 mM and 2.75 mM, respectively. Each response was recorded 3 min after exposure to a new Na^+ level. Transient capacitive currents were roughly compensated by a network with three adjustable time constants. **b**, Sample records of currents evoked by a voltage step from -100 mV to -51 mV, -25 mV, -21 mV, -6 mV and $+59$ mV in Ringer's solution. TTX-insensitive currents were subtracted. **c**, Peak current-voltage relationship by step depolarizations from a holding potential of -100 mV in Ringer's solution. Data from the same oocyte as in **b**.

agreement with a recent report²⁵. It is possible that, because of the presence of microvilli and indentations on the oocyte surface and the resultant large membrane capacitance²³, the entire plasma membrane is not adequately clamped during the generation of the fast inward current²⁶.

The TTX-sensitive fast current induced by the sodium channel II-specific mRNA showed inactivation following a depolarizing prepulse in a time- and voltage-dependent manner. In view of the large membrane capacitance, the steady-state inactivation was measured with a test pulse to 0 mV following a long (2 s) prepulse stepped from -100 mV to different potential levels (-95 to -10 mV). The potential at which activation of the inward current was 50% (ref. 27) ranged from -61 mV to -57 mV. This range was comparable to that observed for the Na^+ current exhibited by oocytes injected with rat brain poly(A)⁺ RNA (-56 to -50 mV) (see also ref. 25).

Our results indicate that the mRNAs derived from the cDNAs encoding the rat brain sodium channel large polypeptides can direct the formation of functional sodium channels in *Xenopus* oocytes, although the magnitude of the current induced by the sodium channel I-specific mRNA is small. The functional properties of the sodium channels resulting from injection of the sodium channel II-specific mRNA are comparable to those of the sodium channels produced in oocytes injected with poly(A)⁺ RNA from rat brain. These findings suggest that the

sodium channel large polypeptide alone is sufficient to exhibit the function of this channel. The possibility that *Xenopus* oocytes themselves contain equivalents to the sodium channel small polypeptides, which may associate with the large polypeptide to form functional sodium channels, cannot be excluded. In this context, we observed that a slowly sedimenting fraction ($<28\text{S}$) of rat brain poly(A)⁺ RNA (presumably containing the mRNAs encoding the sodium channel small polypeptides), when injected into oocytes in combination with the sodium channel II-specific mRNA (or the sodium channel I-specific mRNA with or without the 33-nucleotide insertion⁹), did not enhance the TTX-sensitive response. We also observed that a fast-sedimenting fraction ($\geq 28\text{S}$) of rat brain poly(A)⁺ RNA alone, in contrast with the slowly sedimenting fraction, induced TTX-sensitive inward currents (up to $1.0 \mu\text{A}$) in oocytes, in general agreement with previous reports^{28,29}. The small TTX-sensitive response observed in oocytes injected with the sodium channel I-specific mRNA may be due to inefficient synthesis, modification or transport of the channel protein or to the intrinsic properties of this particular sodium channel.

This investigation was supported in part by research grants from the Ministry of Education, Science and Culture and the Science and Technology Agency of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

Received 21 April; accepted 6 June 1986.

- Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer Associates Inc., Sunderland, Massachusetts, 1984).
- Agnew, W. S., Levinson, S. R., Brabson, J. S. & Raftery, M. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2606–2610 (1978).
- Miller, J. A., Agnew, W. S. & Levinson, S. R. *Biochemistry* **22**, 462–470 (1983).
- Lombet, A. & Lazdunski, M. *Eur. J. Biochem.* **141**, 651–660 (1984).
- Messner, D. J. & Catterall, W. A. *J. biol. Chem.* **260**, 10597–10604 (1985).
- Barchi, R. L. *J. Neurochem.* **40**, 1377–1385 (1983).
- Barchi, R. L., Tanaka, J. C. & Furman, R. E. *J. cell. Biochem.* **26**, 135–146 (1984).
- Noda, M. *et al. Nature* **312**, 121–127 (1984).
- Noda, M. *et al. Nature* **320**, 188–192 (1986).
- Rosenberg, R. L., Tomiko, S. A. & Agnew, W. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1239–1243 (1984).
- Tanaka, J. C., Eccleston, J. F. & Barchi, R. L. *J. biol. Chem.* **258**, 7519–7526 (1983).
- Kraner, S. D., Tanaka, J. C. & Barchi, R. L. *J. biol. Chem.* **260**, 6341–6347 (1985).
- Hanke, W., Boheim, G., Barhanin, J., Pauron, D. & Lazdunski, M. *EMBO J.* **3**, 509–515 (1984).
- Hartshorne, R. P., Keller, B. U., Talvenheimo, J. A., Catterall, W. A. & Montal, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 240–244 (1985).
- Feller, D. J., Talvenheimo, J. A. & Catterall, W. A. *J. biol. Chem.* **260**, 11542–11547 (1985).
- Messner, D. J. & Catterall, W. A. *J. biol. Chem.* **261**, 211–215 (1986).

- Green, M. R., Maniatis, T. & Melton, D. A. *Cell* **32**, 681–694 (1983).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
- Gundersen, C. B., Miledi, R. & Parker, I. *Proc. R. Soc. B220*, 131–140 (1983).
- Gundersen, C. B., Miledi, R. & Parker, I. *Nature* **308**, 421–424 (1984).
- Baud, C., Kado, R. T. & Marcher, K. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3188–3192 (1982).
- Baud, C. & Kado, R. T. *J. Physiol., Lond.* **356**, 275–289 (1984).
- Kusano, K., Miledi, R. & Stinnakre, J. *J. Physiol., Lond.* **328**, 143–170 (1982).
- Barish, M. E. *J. Physiol., Lond.* **342**, 309–325 (1983).
- Leonard, J., Snutch, T., Lubbert, H., Davidson, N. & Lester, H. A. *Biophys. J.* **49**, 386a (1986).
- Methfessel, C. *et al. Pflügers Arch. ges. Physiol.* (in the press).
- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500–544 (1952).
- Sumikawa, K., Parker, I. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7994–7998 (1984).
- Hirono, C. *et al. Brain Res.* **359**, 57–64 (1985).
- Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103–119 (1985).
- Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
- Konarska, M. M., Padgett, R. A. & Sharp, P. A. *Cell* **38**, 731–736 (1984).
- McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838 (1977).
- Miledi, R. & Sumikawa, K. *Biomed. Res.* **3**, 390–399 (1982).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
- Houghton, M. *et al. Nucleic Acids Res.* **8**, 1913–1931 (1980).
- Gurdon, J. B. *The Control of Gene Expression in Animal Development* (Clarendon, Oxford, 1974).
- Sakmann, B. *et al. Nature* **318**, 538–543 (1985).

Analysis of T-cell subsets in rejection of K^b mutant skin allografts differing at class I MHC

Amy S. Rosenberg, Toshiaki Mizuochi & Alfred Singer

Immunology Branch, National Cancer Institute,
National Institutes of Health, Bethesda, Maryland 20892, USA

The T-cell subpopulations which initiate and mediate tissue allograft rejection remain controversial¹⁻⁵. In the present study we attempted to identify the phenotype and function of the T-cell subset(s) primarily responsible for the rejection of skin allografts differing at a single class I locus in the major histocompatibility complex (MHC). We found that the rejection rates by B6 mice ($H-2^b$) of four different class I mutant (K^{bm}) skin allografts form a distinct hierarchy. This hierarchy correlates strikingly and uniquely with the relative precursor frequencies of $Lyt2^+$ interleukin-2-secreting T-helper cells reactive against the various K^{bm} mutants. To investigate the role of $Lyt2^+$ T cells in the rejection of class I-disparate skin allografts directly, $H-2^b$ nude mice were engrafted with K^{bm} skin allografts and then reconstituted with $L3T4^+$ or $Lyt2^+$ T-cell subpopulations from syngeneic $H-2^b$ mice. $Lyt2^+$ T cells were observed to be both necessary and sufficient for the rejection of class I-disparate K^{bm} skin allografts, whereas $L3T4^+$ T cells were neither necessary nor sufficient. These results identify the $Lyt2^+$ interleukin-2-secreting T-cell subset as the critical cell type determining the rejection rate of class I-disparate K^{bm} skin allografts.

Because the cellular mechanisms involved in allograft rejection are likely to be diverse and dependent upon the antigenic disparity expressed by the graft, we limited our present study to the rejection of skin allografts expressing isolated class I MHC disparities. We chose to investigate the rejection of skin allografts from K^b mutant mice^{6,7} by normal B6 mice because (1) each K^b mutant strain in theory differs from wild-type B6 mice only in the expression of a mutant $H-2K^b$ molecule; (2) each K^{bm} class I molecule is well characterized biochemically⁸; (3) the various K^{bm} mutations alter the K^b molecule differently so that they are of varying antigenicity^{9,10}. In the first set of experiments, female B6 mice were engrafted with female bm1, bm3, bm6 or bm10 skin allografts. Each graft was scored daily until rejection or day 150. As can be seen in Fig. 1, the rejection rates form a distinct hierarchy, in which bm1 skin allografts are rejected most rapidly, with a median survival time (MST) of 16 days; bm3 and bm10 skin allografts are rejected less rapidly (MST=22 days) and bm6 skin allografts are rejected least rapidly, the majority of bm6 skin allografts being retained over the 150 days they were followed. A second bm6 skin allograft on mice that failed to reject their initial bm6 graft was also not rejected (data not shown). Nevertheless, the K^{bm6} molecule is not itself defective in triggering skin allograft rejection since bm6 skin allografts are rapidly rejected by control bm1 hosts (MST=17 days) which, unlike normal B6 hosts, can react against both mutant and wild-type epitopes on the K^{bm6} molecule.

We hoped to identify the function and phenotype of the T-cell subset(s) primarily responsible for the observed hierarchy of B6 anti- K^{bm} rejection rates by assessing *in vitro* the precursor frequencies of the various K^{bm} -reactive T-cell subpopulations present in the spleens of normal B6 host animals. We assessed two distinct T-cell functions: (1) cytolytic activity as assayed by the ability to specifically lyse ⁵¹Cr-labelled target cells following stimulation with spleen cells from each K^b mutant strain, and (2) helper activity mediated by interleukin-2 (IL-2) and assayed by the ability to secrete IL-2 in response to stimulation with mutant spleen cells¹¹. Fractionated $L3T4^+$ or $Lyt2^+$ T-cell sub-

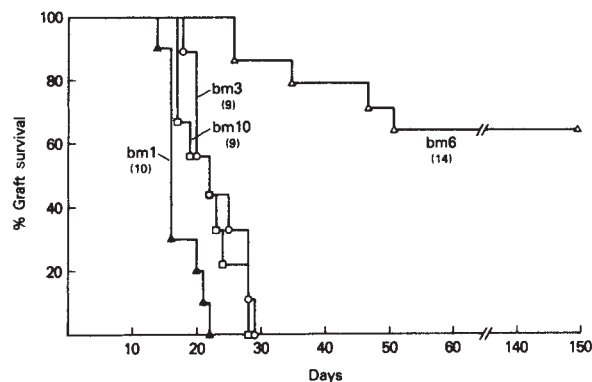


Fig. 1 Rejection of K^{bm} skin allografts by B6 mice. Female B6 mice were engrafted with a tail-skin allograft from female bm1 ($\blacktriangle$), bm3 ($\circ$), bm6 ($\triangle$) or bm10 ($\square$) mice using an adaptation of the method of Billingham and Medawar²⁹. Grafts were scored daily until rejection or day 150. Numbers in parentheses indicate the number of mice in each group. Control B6 grafts were not rejected.

populations were obtained from the spleens of normal B6 mice by pretreatment with either anti- $Lyt2.2$ or anti- $L3T4$ monoclonal antibody and complement. The average precursor frequency determinations from several limiting dilution experiments are shown in Table 1, together with the skin graft rejection rates determined from Fig. 1. The precursor frequencies of anti- K^{bm} $L3T4^+$ cytolytic T lymphocytes (T_c) are not given in Table 1 because none were detected¹². In contrast, anti- K^{bm} $L3T4^+$ IL-2-secreting helper (T_h) cells were detected, consistent with previous observations that $L3T4^+$ T cells recognize class I alloantigens in the context of self-Ia determinants and function as IL-2-secreting T_h cells but not as T_c effectors in class I allospecific responses¹³⁻¹⁵. It can be seen in Table 1 that bm1-reactive T cells were the most frequent of any of the anti- K^{bm} specificities examined, regardless of the function or phenotype of the T-cell subset assessed. This result is consistent with the observation that bm1 skin allografts were the most rapidly rejected, but is unhelpful in identifying the T-cell subset(s) primarily responsible for determining allograft rejection rate. It can also be seen in Table 1 that the precursor frequencies of anti-bm6 cells were approximately equal to the precursor frequencies of anti-bm3 and anti-bm10 cells in both unseparated and $L3T4^+$ T_h -cell populations, as well as in both unseparated and $Lyt2^+$ T_c populations. These results are markedly discordant with the slow and variable rejection of bm6 versus bm3 or bm10 skin allografts, and suggest that none of these T-cell subsets is the one primarily responsible for determining the rejection rate of class I-disparate K^{bm} skin allografts. However, most interestingly, it can be seen in Table 1 that the relative precursor frequencies of K^{bm} -reactive T cells in the $Lyt2^+$ IL-2-secreting T_h -cell subset correlate perfectly with the rates at which skin allografts from these various mutant strains are rejected. Most B6 mice fail to reject bm6 skin allografts and B6 spleen-cell populations contain few, if any, $Lyt2^+$ T cells that secrete IL-2 in response to bm6 stimulator cells. The K^{bm6} molecule effectively triggers $Lyt2^+$ T cells from $H-2K^k$ mice to secrete IL-2¹⁶, so that the failure to detect significant numbers of bm6-reactive $Lyt2^+$ IL-2 secreting T cells in B6 mice reflects their low frequency and is not the result of a unique defect in the ability of the K^{bm6} molecule to trigger $Lyt2^+$ T cells to secrete IL-2. These data imply that the number of specifically reactive IL-2-secreting $Lyt2^+$ T cells is the limiting and critical determinant of the rate at which rejection of K^{bm} skin allografts can occur and suggest that $Lyt2^+$ T cells in general, and IL-2-secreting $Lyt2^+$ T cells in particular, are required for such allografts to be rejected at all.

To examine the role of $Lyt2^+$ T cells in the rejection of class I-disparate K^{bm} skin allografts directly, we made use of an experimental model in which isolated T-cell subpopulations can

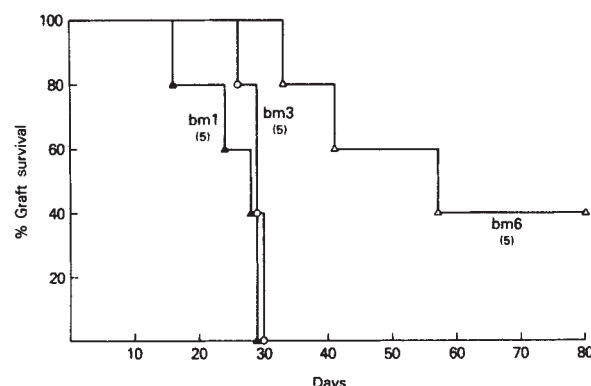


Fig. 2 Rejection of K^{bm} skin allografts by reconstituted B10 nude mice. Female B10 nude mice were engrafted on day 0 with tail-skin allografts from female bm1 ($\blacktriangle$), bm3 ($\circ$) or bm6 ($\triangle$) mice. On day 1, each mouse was injected intravenously with 50×10^6 untreated spleen cells from unprimed B6 female mice. Grafts were scored daily until rejection or day 80. Numbers in parentheses indicate the number of mice in each group. MST in days were as follows: bm1 = 28, bm3 = 29, bm6 = 57. Control B6 grafts were not rejected.

be assessed for their ability to mediate skin allograft rejection after transfer into T-cell-deficient *nu/nu* mice¹⁷. H-2^b nude mice were engrafted with skin allografts on day 0, reconstituted with 50×10^6 H-2^b spleen cells or spleen cell subpopulations on day 1, and then examined daily for the status of their skin grafts. To ascertain that skin graft rejection in this experimental model depended upon the adoptive transfer of immunocompetent T cells, we engrafted B10 nude mice with B10.BR (H-2^k) skin and reconstituted them with 50×10^6 spleen cells from either normal or T-cell-deficient, nude B10 mice. The H-2^k skin allografts were rapidly and completely rejected by nude mice reconstituted with normal B10 spleen cells (MST = 14 days) but were not rejected by any of the nude mice reconstituted with T-cell-deficient, nude B10 spleen cells (MST > 33 days). To determine if the mechanisms of skin graft rejection in this experimental model reflect those operative in normal mice, we engrafted B10 nude mice with bm1, bm3 or bm6 skin allografts and reconstituted them with normal B6 spleen cells (Fig. 2). Precisely the same hierarchy of rejection rates was observed as had been seen in normal mice, namely $bm1 > bm3 > bm6$, suggesting that similar rejection mechanisms are involved (Fig. 2). Finally, to identify the phenotype of the T cells able to effect the rejection of class I-disparate K^{bm} allografts, we engrafted B10 nude mice with bm1 skin allografts and then reconstituted them with 50×10^6 untreated normal B6 spleen cells or 50×10^6 normal B6

spleen cells that had been depleted either of $Lyt2^+$ cells or of $L3T4^+$ cells (Fig. 3). The depletion procedures were >99% effective as shown by immunofluorescence and flow microfluorometry (data not shown). The results (Fig. 3) demonstrated that the isolated $Lyt2^+$ T-cell population was both necessary and sufficient to reject bm1 skin allografts. Interestingly, nude mice reconstituted with the isolated $Lyt2^+$ T-cell population rejected bm1 skin allografts more rapidly (MST = 10 days) than did nude mice reconstituted with unfractionated T-cell populations containing both $Lyt2^+$ and $L3T4^+$ T cells (MST = 22 days), indicating that $L3T4^+$ T cells do not synergize with $Lyt2^+$ T cells to effect the rejection of bm1 skin allografts and that they might conceivably interfere with the rejection. Finally, nude mice reconstituted with the isolated $L3T4^+$ T-cell population were incapable of rejecting bm1 skin allografts (MST > 60 days) (Fig. 3), although they were immunocompetent as indicated by their rapid rejection of control class II-disparate skin grafts (data not shown).

The present study provides strong support for the role of $Lyt2^+$ T cells in general, and IL-2-secreting $Lyt2^+$ T cells in particular, in skin allograft rejection. It significantly extends earlier observations¹⁸⁻²² by demonstrating that $Lyt2^+$ T cells are not only able to mediate allograft rejection but are the predominant T cells in the rejection of class I-disparate K^{bm} skin allografts. Interestingly, the presence in B6 mice of both K^{bm} .

Table 1 Rejection rates of K^{bm} skin allografts and precursor frequencies of K^{bm} -reactive T cell subsets

Responder	Strain	Target	Graft survival MST (days)*	Precursor frequencies ($\times 10^3$) of K^{bm} -specific T cells in unprimed B6 spleen-cell populations				
				Unseparated	IL-2-secreting T cells†	Cytolytic T cells‡		
B6		bm1	16	1/16	1/44	1/24	1/3	1/0.6
B6		bm3	22	1/61	1/136	1/73	1/6	1/5
B6		bm10	22	1/53	1/45	1/93	1/23	1/6
B6		bm5	>150	1/59	1/67	<1/1,000	1/6	1/4

* Obtained from Fig. 1.

† Summary of seven experiments. Each limiting-dilution titration consisted of 6-8 experimental points, representing 24-48 replicate cultures. Individual 0.2-ml cultures contained graded numbers of untreated or fractionated responder spleen-cell populations, 5×10^5 2,000R-irradiated stimulator spleen cells and 0.01% (v/v) ascites of 7D4 monoclonal antibody²⁵, specific for the murine IL-2 receptor. The addition of monoclonal anti-IL-2 receptor antibody to the response cultures significantly improves the accuracy of IL-2 production measurements by blocking the consumption of IL-2 during the culture period without interfering with either the production of IL-2 or its subsequent measurement¹¹. After 5 days of culture at 37°C and 5% CO₂, 0.1 ml of supernatant was collected from each well and assayed for their ability to maintain the proliferation of 4×10^3 indicator HT-2 cells²⁶. Individual cultures were considered positive if the ³H-thymidine incorporation by HT-2 cells in response to their supernatant was more than 3 s.d. above that stimulated by medium alone. Minimal estimates of the frequencies of T_H cells were calculated by analysis of the Poisson distribution relationship between the percentage of negative cultures per group and the number of responding cells²⁷. The T-cell fractionation procedures using monoclonal anti-L3T4 and anti-Lyt2 antibodies are described in Fig. 3 legend.

‡ Summary of two experiments. Each limiting-dilution titration consisted of 6-8 experimental points, representing 24-48 replicate cultures. Individual 0.2-ml cultures contained graded numbers of untreated or fractionated responder spleen-cell populations, 5×10^5 2,000R-irradiated stimulator spleen cells, and 12.5% supernatant from concanavalin A-induced spleen-cell cultures²⁸, to which α -methyl mannoside had been added. After 7 days of culture at 37°C and 7.5% CO₂, 4×10^3 ⁵¹Cr-labelled concanavalin A-induced spleen blast cells were added to each culture as target cells in a 4-h ⁵¹Cr-release assay. Cultures were considered positive if their ⁵¹Cr release was 3 s.d. above spontaneous release. Minimal estimates of the frequencies of T_C precursors were calculated as previously described²⁷. The T-cell fractionation procedures are described in Fig. 3 legend.

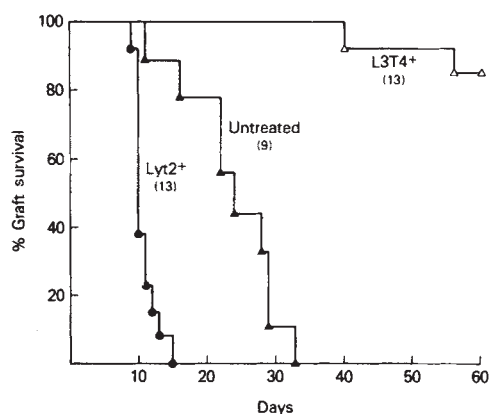


Fig. 3 Ability of isolated T-cell subpopulations to reject bm1 skin allografts. Female B10 nude mice were engrafted on day 0 with tail-skin allografts from female bm1 mice. On day 1, the mice were injected intravenously with 50×10^6 B6 spleen cells which: were untreated and so contained both L3T4⁺ and Lyt2⁺ T-cell subpopulations (Δ); had been pretreated with anti-Lyt2.2 monoclonal antibody + complement and so contained only the L3T4⁺ T-cell subset ($\triangle$); had been pretreated with anti-L3T4 monoclonal antibody + complement and so contained only the Lyt2⁺ T-cell subset ($\bullet$). Numbers in parentheses indicate the number of mice in each group. Grafts were scored daily until rejection or day 60. MST in days of bm1 skin allografts on B10 nude mice reconstituted with various T-cell populations were as follows: Lyt2⁺ T cells = 10; unfractionated T cells = 22; L3T4⁺ T cells > 60.

Methods. Anti-L3T4 monoclonal antibody was a culture supernatant of the hybridoma cell line GK1.5 (ref. 30) provided by Dr Frank Fitch. Anti-Lyt2.2 monoclonal antibody was a culture supernatant of the hybridoma cell line 83-12-5 provided by Dr Jeffrey Bluestone. Depletion of L3T4⁺ T cells or Lyt2⁺ cells was accomplished by incubating spleen cells at a density of 10^7 cells ml⁻¹ with anti-L3T4 (1:2 dilution of culture supernatant) or anti-Lyt2.2 monoclonal antibody (1:5 dilution of culture supernatant) for 30 min on ice. Cells were then pelleted, resuspended and incubated with complement for 40 min at 37 °C. Treated cells were washed three times before injection into experimental animals or before addition to experimental cultures. Efficacy of each treatment (>99%) was confirmed by immunofluorescence and flow microfluorimetry. Spleen cells from each group of engrafted and adoptively reconstituted nude mice were assessed 30 days after reconstitution by immunofluorescence and flow microfluorimetry for numbers of L3T4⁺ and Lyt2⁺ T cells. Spleens from bm1 engrafted nude mice reconstituted 30 days previously with L3T4⁺ T-cell populations contained <1% Lyt2⁺ T cells, whereas spleens from bm1 engrafted nude mice reconstituted 30 days previously with Lyt2⁺ T-cell populations contained <1% L3T4⁺ cells.

reactive L3T4⁺ T_H and Lyt2⁺ T_C is not predictive of whether or not K^{bm} skin is rejected, even though these two T-cell subpopulations collaborate effectively *in vitro* to generate K^{bm}-specific T_C responses^{13,15,23}. Since class I-reactive L3T4⁺ T_H cells recognize a composite determinant composed of class I + class II MHC molecules¹³⁻¹⁵, the deficient expression of class I MHC on class II-bearing skin epidermal cells²⁴ provides a plausible explanation for the failure of these T_H cells to participate in the rejection of K^{bm} allografts. Thus, class I-reactive L3T4⁺ T_H cells may never encounter a class (I + II) antigenic complex on individual cells within a skin allograft.

In conclusion, the present study makes three points regarding the rejection of class I-disparate K^{bm} skin allografts by wild-type H-2^b mice: first, Lyt2⁺ T cells are the only T cells required to effect the rejection of class I-disparate K^{bm} skin allografts; second, while Lyt2⁺ T_C effectors may contribute importantly to the rejection of K^{bm} skin allografts, it is the number of IL-2-secreting, K^{bm}-reactive T cells within the Lyt2⁺ T-cell subpopulation that is limiting and that determines the rates at which K^{bm} skin grafts are rejected; and third, potential collaborations

between anti-K^{bm} L3T4⁺ T_H cells and Lyt2⁺ T_C precursors do not have a prominent role in the rejection of K^{bm} skin allografts.

We thank Dr Jeffrey Bluestone for helpful discussions and Drs Ronald Gress, Richard Hodes, Susan McCarthy, David Sachs and Dinah Singer for critical readings of the manuscript.

Received 13 March; accepted 8 May 1986.

- Loveland, B. E., Hogarth, P. M., Ceredig, R. H. & McKenzie, I. F. C. *J. exp. Med.* **153**, 1044-1057 (1981).
- LeFrancis, L. & Bevan, M. J. *J. exp. Med.* **159**, 57-67 (1984).
- Hall, B. M. & Dorsch, S. E. *Immun. Rev.* **77**, 31-59 (1984).
- Lafferty, K. J. & Simeonovic, C. J. *Transplant. Proc.* **16**, 927-930 (1984).
- Steinmuller, D. *Transplantation* **40**, 229-233 (1985).
- Melvold, R. W. & Kohn, H. I. *Immunogenetics* **3**, 185-191 (1976).
- Kohn, H. I. *et al. Immunogenetics* **77**, 279-294 (1978).
- Nairn, R., Yamaga, K. & Nathenson, S. G. *A. Rev. Genet.* **14**, 241-277 (1980).
- Melief, C. J. M., Schwartz, R. S., Kohn, H. I. & Melvold, R. W. *Immunogenetics* **2**, 337-348 (1975).
- Melief, C. J. M., de Waal, L. P., van der Meulen, M. Y., Melvold, R. W. & Kohn, H. I. *J. exp. Med.* **151**, 993-1013 (1980).
- Mizuuchi, T., Ono, S., Malek, T. R. & Singer, A. *J. exp. Med.* **163**, 603-619 (1986).
- Golding, H. & Singer, A. *J. Immun.* **135**, 1610-1615 (1985).
- Singer, A., Kruisbeek, A. M. & Andrysiak, P. M. *J. Immun.* **132**, 2199-2209 (1984).
- Golding, H. & Singer, A. *J. Immun.* **133**, 597-605 (1984).
- Mizuuchi, T. *et al. J. exp. Med.* **162**, 427-433 (1985).
- Mizuuchi, T. *et al. J. Immun.* (in the press).
- Lynch, D. H., Weiland, D. J., Rosenberg, S. A. & Hodes, R. J., *Transplantation* (in the press).
- Engers, H. D., Galsebrook, A. L. & Sorenson, G. D. *J. exp. Med.* **156**, 1280-1285 (1982).
- Engers, H. D., Sorenson, G. S., Geronimo, T., Horvath, C. & Brunner, K. T. *J. Immun.* **129**, 1292-1298 (1982).
- Mizuuchi, T., Agostino, M. & Lafferty, K. J. *Transpl. Proc.* **17**, 1644-1645 (1985).
- Warren, H. S., Prouse, S. J., Agostino, M. & Lafferty, K. J. *Transpl. Proc.* **16**, 954-955 (1984).
- Sprent, J., Schaefer, M., Lo, D. & Korngold, R. *J. exp. Med.* **163**, 998-1011 (1986).
- Weinberger, O., Germain, R. N. & Burakoff, S. J. *Nature* **302**, 429-431 (1983).
- Caughman, S. W. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Malek, T. R., Robb, R. J. & Shevach, E. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5694-5698 (1983).
- Watson, J. *J. exp. Med.* **150**, 1510-1519 (1979).
- Taswell, C. *J. Immun.* **126**, 1614-1619 (1981).
- Kruisbeek, A. M., Zystra, J. J. & Kroese, T. J. M. *J. Immun.* **125**, 995-1002 (1980).
- Billingham, R. E. & Medawar, P. B. *J. exp. Biol.* **28**, 385-401 (1951).
- Dialynas, D. P. *et al. Immun. Rev.* **74**, 29-56 (1983).

Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytotoxicity

Jürg Tschopp*, Danièle Masson* & Keith K. Stanley†

* Institute of Biochemistry, University of Lausanne, CH-1066 Epalinges, Switzerland
† EMBL, Meyerhofstrasse 1, Postfach 10.2209, D-6900 Heidelberg, FRG

Cytotoxicity mediated by complement or cytotoxic lymphocytes results in the formation of morphology similar lesions in the target membrane¹⁻⁶. These lesions, formed by the polymerization of C9 or perforin respectively, contribute the major killing action by causing osmotic lysis of the target cell. Following the suggestion of Mayer⁷ that the mechanisms of humoral and cell-mediated cytotoxicity might be related, studies into the morphology of the membrane lesions formed⁸⁻¹³, and the proteins responsible for causing the lesions^{9,14,15}, have shown several similarities. While the lesion caused by natural and T-killer cells is a little larger than that caused by complement, its overall shape is similar and in both cases the cylindrical pore is formed by polymerization of a monomeric subunit, C9 (relative molecular mass, M_r = 71,000) for complement⁹, and perforin (M_r = 66,000) for cell-mediated cytotoxicity^{14,15}. C9 has an absolute requirement for a receptor in the target membrane formed by the earlier membrane attack complex components, C5b, C6, C7 and C8 (ref. 8). For perforin, polymerization in a target membrane requires no receptor, specificity being derived from the specific recognition between killer and target cell. Both proteins can be made to polymerize *in vitro* by the addition of divalent cations (Zn^{2+} for C9 (ref. 16) and Ca^{2+} for perforin^{10,12,13}) and the resultant complexes closely resemble their physiological counterparts. Antibodies raised

against lymphocyte-killed targets have also been shown to cross-react with complement proteins^{17,18,33}, but the antigenically related proteins were not determined in these studies. We show here using purified proteins that perforin, C9 and complexes involving C7 and C8 share a common antigenic determinant which is probably involved in polymerization.

During the expression screening of a human cDNA library in pEX (ref. 19) it was noted that one clone, which did not contain C9 sequences, reacted strongly with a polyclonal antibody of high specificity to C9 (clone C9.1 in ref. 19). When the antibody affinity purified on this fusion protein was tested in Western blots of purified proteins it was found that it bound strongly to mouse perforin. This fusion protein was therefore used to boost rabbits initially injected with mouse perforin.

Figure 1 shows the Western blot analysis using this particular antiserum named C9.1 (Fig. 1a). Not only were purified perforin and C9 immunostained, but a single band appeared with the same apparent M_r as perforin when whole cytoplasmic cytotoxic T-lymphocyte (CTL) granules were analysed. The antiserum did not react with bovine serum albumin or other serum proteins; however, in half of the experiments, C7 was also detected (data not shown).

When a panel of monoclonal antibodies raised against human monomeric C9 was tested, none reacted with mouse perforin, but one monoclonal antibody (BC5), originally described as binding only to epitopes present on poly C9 (ref. 19), did cross-react. This neoantigen-specific antibody binds to a non-consecutive epitope, and thus binds only weakly to C9 separated in the presence of SDS and transferred to nitrocellulose. In contrast, dot-blot analysis, which presumably conserves this epitope, can be used for its analysis (Fig. 1b). Less than 160 pg of poly-C9 can be immunostained using this antibody, whereas C5b-6, C7, C8 and monomeric C9 do not stain. In contrast, when the late components are first mixed to form the respective complex and then applied to the nitrocellulose, C5b-7, C5b-8, C5b-9 (comprising poly C9) and poly-C9 are all stained. These complexes must therefore exhibit a common epitope not present on the individual precursor proteins.

The amino-acid sequence of C9^{19,20} contains two cysteine-rich regions¹⁹ which show sequence and predicted structural homology²¹ to the low-density lipoprotein (LDL) receptor and urokinase. The function of these regions is unknown, but one hypothesis is that these regions form compact globular domains involved in the polymerization of C9. We therefore postulated that if mouse perforin and human C9 shared a common domain, it was most likely to be found in one of these conserved regions, since both molecules polymerize in a similar manner.

Two peptides were synthesized corresponding to amino acids 74-86 (peptide A) and 101-111 (peptide B) of C9, that is, the segment with high homology to the LDL receptor (Fig. 2). Antibodies raised against the peptides coupled to ovalbumin were tested for their reactivity to C9 and perforin. Both anti-peptide A and anti-peptide B immunostained human C9 and mouse perforin (Fig. 3). Perforin staining was considerably weaker, especially when immunoblotting was performed using anti-peptide A. These results suggest that the amino-acid sequence of perforin must be slightly different, especially in segment 74-86 of C9. This second segment also shows a lower degree of homology to the LDL receptor than does the segment corresponding to peptide B. In addition to C9, anti-peptide B bound to C8 α (ref. 22), to C7, and weakly to C6. Anti-peptide A also reacted with C8 α , C7 and C6, although binding to C7 was extremely weak. Other proteins tested, such as bovine serum albumin, were not stained at all with anti-peptide B²² and anti-peptide A (Fig. 3). Binding could also be inhibited by addition of the corresponding peptide. This was true for anti-body binding to C9 (Fig. 3), C8 α (ref. 22), C7, C6 and perforin (data not shown).

To exclude the possibility that the binding of the anti-peptide antibodies to these proteins was a chance event, the Dayhoff

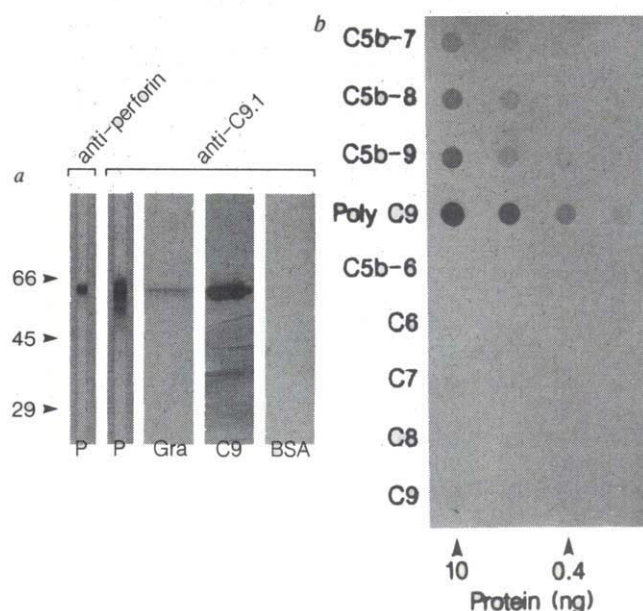


Fig. 1 *a*, Cross-reactivity of C9 with perforin. Perforin (P), granules (Gra), C9 and bovine serum albumin (BSA) were separated under reducing conditions on a 10% polyacrylamide²⁷ gel in the presence of SDS (SDS-PAGE) and transferred to nitrocellulose²⁸. As a control, perforin was incubated with a specific rabbit anti-perforin antiserum under the same conditions (left panel). *b*, Dot-blot analysis of complexes of the late components of complement using monoclonal antibody BC5 as a probe.

Methods. *a*, The nitrocellulose membrane was incubated for 2 h in a buffer containing 20 mM Tris-HCl pH 8.4, 150 mM NaCl, 5 mM EDTA, 1% gelatin and 0.1% BSA. Anti-C9.1 antiserum was then added at a 1:100 dilution in the same buffer. The blot was developed with protein-A-peroxidase conjugate using 4-chloro-1-naphthol and H₂O₂ as substrates. Perforin and granules were isolated from the mouse CTL line B6.1 as described elsewhere^{12,14}. Antibody C9.1 was raised in rabbits by injection of 20 μ g purified mouse perforin in Freund's complete adjuvant into popliteal lymph nodes of rabbits followed by three boosts with 20 μ g of perforin administered intradermally. Fusion protein from clone C9.1 (ref. 19) was then injected intradermally three times at intervals of 3 weeks and the animal bled 11 days after the final injection. *b*, Complement proteins C5b-6, C7, C8 and C9 were prepared according to standard procedures²⁹. 10 ng, 2 ng, 0.4 ng, 80 pg of the proteins were adsorbed to nitrocellulose using the dot-blot microfiltration apparatus (BioRad). The different complexes were formed by mixing the respective proteins C5b-6 (1 μ g), C7 (0.5 μ g), C8 (0.6 μ g) and C9 (1.3 μ g) in 100 μ l Tris-buffered saline (TBS, 10 mM Tris-HCl pH 7.4 containing 150 mM NaCl), incubating them at 20°C for 15 min (in the case of the C5b-9 complex, incubation temperature was 37°C). Poly-C9 was formed by incubating 10 μ g C9 in 10 μ l TBS containing 10⁻⁴ M ZnCl₂ for 2 h at 37°C. The protein mixtures were diluted to the appropriate concentration and then applied onto the nitrocellulose membrane. Immunostaining was performed as described in *a*, using a 1/500 dilution of the monoclonal antibody.

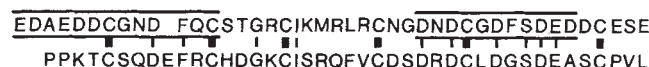


Fig. 2 Comparison of the homology regions of C9 and the LDL receptor^{19,24}. The peptides synthesized are indicated in the boxed area.

Methods. Peptides were synthesized according to a modified Merrifield method³⁰, using (fluoroenylmethoxycarbonyl) protected amino acids (Bachem). A semiautomatic peptide synthesizer (Labor-tec) was used for the synthesis. Peptides were purified on a preparative Mono Q column (Pharmacia, FPLC system).

Fig. 3 Western blot analysis of C6, C7, C8, C9 and perforin (P). Samples of 3 μ g of C6 (M_r 115,000), C7 (M_r 110,000), C8 (M_r 68,000, C8 α , M_r 65,000, C8 β , M_r 15,000), C9 (M_r 71,000) and perforin (M_r 66,000) were separated under reducing conditions by SDS-PAGE and transferred to nitrocellulose²⁸. Anti-peptide A antiserum and anti-peptide B antisera were added at a 1:100 dilution. Binding of the antibodies was revealed using protein-A-peroxidase conjugates (1:1,000 dilution, Sigma). The specificity of anti-peptide A antiserum is shown in the right panel: nitrocellulose-bound C9 was incubated with the antiserum in the presence and absence of 100 μ M peptide A. Blotted BSA incubated with anti-peptide A is shown as an additional control. Antisera to the peptides were raised by injecting 200 μ g of peptides coupled to ovalbumin with *bis*-diazobenzidine³¹ in complete Freund's adjuvant on day 0. The animals were boosted with the same quantity of peptide in incomplete Freund's adjuvant on days 14 and 28 and were bled on day 35.

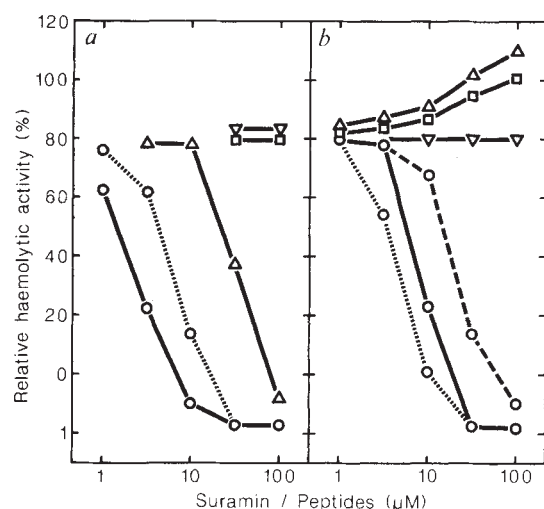
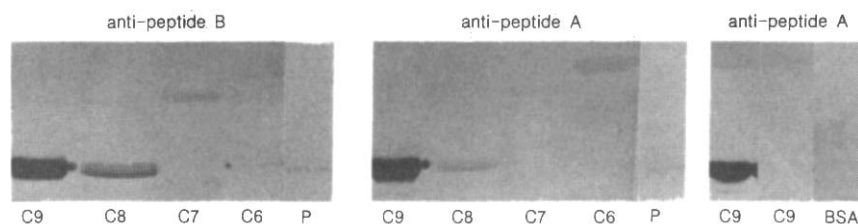


Fig. 4 Effect of peptides and suramin on C7, C8, C9, CTL granule and perforin haemolytic activity. *a*, Inhibition of perforin (dashed line) and granule activity (solid line) by peptide A (□), peptide B (Δ), a peptide (CDDPPPEIPHAT, ▽) derived from the sequence of the interleukin-2 receptor³² and suramin (○) (Bayer, Wuppertal). *b*, Inhibition of C7 (dashed line), C8 (dotted line) and C9 activity (solid line) with the agents detailed in *a*.

Methods. Granules (100 μ l) isolated from the CTL line B6.1 (ref. 12) or purified perforin¹⁴ were adjusted to a concentration lysing 80% of the sheep red blood cells, and mixed with the peptides/suramin and 300 μ l sheep erythrocytes (1.5×10^7 cells per ml) in a solution containing 10 mM Veronal buffer pH 7.4, 142 mM NaCl, 0.1% gelatin, 1 mM Ca^{2+} and 0.15 mM Mg^{2+} (GVB). The mixture was incubated for 15 min at 37 °C, the erythrocytes spun down at 1,500 r.p.m. for 5 min, and the degree of haemolysis measured by determining the haemoglobin release at 412 nm in a Gilford spectrophotometer. The haemolytic activity was calculated as *Z* number¹³. The decrease in the *Z* number was then transformed into relative haemolytic activity. *b*, C9 haemolytic test⁹ was performed by diluting C9 to a concentration causing 80% lysis of the erythrocytes (usually 1–2 ng ml^{-1}). C9, inhibitors and erythrocytes bearing complement components C1–C8 on their surface (1.5×10^7 cells per ml) were then mixed in a total volume of 400 μ l GVB and incubated for 30 min at 37 °C. The degree of lysis was determined as outlined above. C8 activity²² was carried out by incubating erythrocytes having C1 to C7 bound (1.5×10^7 cells per ml) with inhibitors and 0.1 $\mu\text{g ml}^{-1}$ C8 for 15 min at 37 °C in 400 μ l GVB. The cells were washed twice in GVB and resuspended in 400 μ l GVB containing 1 $\mu\text{g ml}^{-1}$ C9. Incubation was continued for 30 min at 37 °C and the degree of haemoglobin release determined. C7 activity was determined accordingly with C1–C6 bearing erythrocytes. The inhibitors were added with C7 (0.1 $\mu\text{g ml}^{-1}$) and after washing, excess of C8 and C9 was added. C6 activity could not be determined, since erythrocytes carrying C1–C5b lose C6 binding activity instantaneously.

protein sequence data bank was scrutinized for the presence of one of the sequences corresponding to the C9-derived peptides. Using the MATCH routine of the Protein Identification Resource²³, we searched for proteins having 6 out of the 11–13 amino acids of peptide A or peptide B in common on the assumption that 6 amino acids are sufficient for an antigenic site. Seven proteins (including the LDL receptor) fulfilled this criterion for peptide B and 15 proteins for peptide A. However, only C9 showed up in both searches, indicating the low probability that a random protein could be detected by the two anti-peptide antibodies.

The above results thus suggest that C6, C7, C8 α , C9 and perforin all contain structures homologous to the amino-terminal cysteine-rich region of the LDL receptor. In the LDL receptor, this region binds to the LDL apoprotein B²⁴ in a manner inhibited by two negatively charged molecules, suramin²⁵ and heparin²⁶, suggesting an electrostatic interaction with its ligand. Since a similar interaction within the C9 molecule might account for its polymerization, we tested suramin and the peptides A and B for a direct effect on cytolysis. Figure 4 shows that both C9 and perforin activity were completely abrogated at concentrations above 30 μM and 10 μM suramin, respectively. A twofold higher concentration was required for the suppression of perforin activity when perforin was present in CTL granules. C7 and C8 activity were suppressed over a similar range of concentrations.

When the peptides themselves were tested for their inhibitory action, peptide B was found to inhibit granule-mediated cytolysis at concentrations above 100 μM , but C9 activity was unexpectedly enhanced. At 100 μM peptide concentration, C9 activity was augmented 1.3-fold. The same enhancing effect on C9 activity was observed when peptide A was tested, although this peptide did not affect granule or perforin activity (Fig. 4). As a control, a peptide derived from the interleukin-2 receptor showed no effect at all.

We have shown that C9 and perforin contain one or more similar antigenic determinants which cross-react with antibodies raised against intact proteins and synthetic peptides. One of these similar antigenic regions is located in the segment of C9 bearing homology to the LDL receptor. The ability of suramin to inhibit cytolytic activity of C9 suggests that the negatively charged region of the LDL receptor homology region is involved in C9 polymerization. Peptide B, which encompasses this negatively charged region, has opposite effects on perforin and C9 activity. This is particularly interesting because C5b8 complexes, which are essential components in a target membrane for C9 activity, also contain a determinant antigenically related to peptide B. It is possible therefore that the potentiation of C9 activity in the presence of peptide B is also a reflection of the catalytic activity of C5b8.

We thank Dr H. Haymerle and Dr E. Mollnes for antisera, E. Burnier, R. Etges and Z. Freiwald for help in preparing the manuscript, and Ch. Estrade and S. Schäfer for technical assistance. J.T. is supported by a grant of the Swiss NSF.

Received 25 February; accepted 22 April 1986.

1. Bhakdi, S. & Tranum-Jensen, J. *Biochim. biophys. Acta* **737**, 343-372 (1983).
2. Lachmann, P. J. & Thompson, R. A. *J. exp. Med.* **131**, 643-652 (1970).
3. Müller-Eberhard, H. J. *A. Rev. Biochem.* **44**, 697-723 (1975).
4. Lachmann, P. J. *Nature* **305**, 473 (1983).
5. Podack, E. R. *Immun. Today* **6**, 21-27 (1985).
6. Dourmashkin, R. R., Deteix, P., Simon, C. B. & Henkart, P. *Clin. exp. Immun.* **42**, 554-559 (1980).
7. Mayer, M. M. *J. Immun.* **119**, 1195-1202 (1977).
8. Tschopp, J., Müller-Eberhard, H. J. & Podack, E. R. *Nature* **298**, 534-538 (1982).
9. Podack, E. R. & Tschopp, J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 574-578 (1982).
10. Henkart, P. A., Millard, P. J., Reynolds, C. W. & Henkart, M. P. *J. exp. Med.* **160**, 75-91 (1984).
11. Dennert, G. & Podack, E. R. *J. exp. Med.* **157**, 1483-1495 (1983).
12. Masson, D., Corthésy, P., Nabholz, M. & Tschopp, J. *EMBO J.* **4**, 2533-2538 (1985).
13. Podack, E. R. & Königsberg, P. J. *J. exp. Med.* **160**, 695-710 (1984).
14. Masson, D. & Tschopp, J. *J. biol. Chem.* **260**, 9096-9072 (1985).
15. Podack, E. R., Young, J. D. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8629-8633 (1985).
16. Tschopp, J. *J. biol. Chem.* **259**, 10569-10573 (1984).
17. Sundsmo, J. S. & Müller-Eberhard, H. J. *J. Immun.* **122**, 2371-2378 (1975).
18. Ward, R. H. R. & Lachmann, P. J. *Immunology* **56**, 179-188 (1985).
19. Stanley, K. K., Kocher, H. P., Luzio, J. P., Jackson, P. & Tschopp, J. *EMBO J.* **4**, 375-382 (1985).
20. DiScipio, R. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 7298-7302 (1984).
21. Stanley, K. K., Page, M., Campbell, A. K. & Luzio, J. P. *Molec. Immun.* (in the press).
22. Tschopp, J. & Möllnes, T. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4223-4227 (1986).
23. George, D. G., Barker, W. C. & Hunt, L. T. *Nucleic Acids Res.* **14**, 11-15 (1986).
24. Südhof, T. C., Goldstein, J. L., Brown, M. S. & Russel, D. W. *Science* **228**, 815-822 (1985).
25. Schneider, W. J., Beisiegel, U., Goldstein, J. L. & Brown, M. S. *J. biol. Chem.* **257**, 2664-2673 (1982).
26. Goldstein, J. L., Basu, S. K., Brunschede, G. Y. & Brown, M. S. *Cell* **7**, 85-95 (1976).
27. Lämml, U. K. *Nature* **227**, 680-685 (1970).
28. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4358 (1979).
29. Podack, E. R. & Tschopp, J. *Molec. Immun.* **21**, 589-603 (1984).
30. Atherton, E., Gait, M. J., Sheppard, R. C. & Williams, B. J. *Bioorg. Chem.* **8**, 351-370 (1979).
31. Bassiri, R. M., Dvorak, J. & Utiger, R. D. in *Methods of Hormone Radioimmunoassay* (eds Jaffe, B. M. & Behrman, H. R.) 46-47 (Academic, New York, 1979).
32. Cosman, D. *et al. Nature* **312**, 768-771 (1984).
33. Zalman *et al. Bioscience rep.* **5**, 1093-1095 (1985).

Probing the phagolysosomal environment of human macrophages with a Ca^{2+} -responsive operon fusion in *Yersinia pestis*

Charles Pollack*, Susan C. Straley† & Mark S. Klempner*‡

* Department of Medicine, Tufts University School of Medicine, New England Medical Center, Boston, Massachusetts 02111, USA

† Department of Medical Microbiology and Immunology, Albert B. Chandler Medical Center, University of Kentucky, Lexington, Kentucky 40536, USA

Several microorganisms, including *Yersinia* sp., *Salmonella* sp., *Brucella* sp., *Mycobacterium* sp. and *Leishmania* sp., have successfully adapted to grow within macrophage phagolysosomes. Infections caused by these intracellular pathogens are among the most difficult to treat. As part of an antimicrobial strategy directed at modifying the phagolysosomal environment to the disadvantage of these important pathogens, we are defining the ambient conditions within the organism-containing phagolysosome. To probe this environment, we have used *Yersinia pestis*, whose expression of several virulence attributes is highly dependent on the Ca^{2+} concentration in its growth environment. We first genetically engineered a strain of *Y. pestis* which responds to a low-calcium environment by transcription of inserted structural genes of the *Escherichia coli* *lac* operon. Using this mutant organism as a relevant biological probe, we demonstrate here that the calcium concentration in *Y. pestis*-containing phagolysosomes is sufficiently low to permit virulence gene expression; this resolves the question of where *Y. pestis* might express its Ca^{2+} -regulated genes *in vivo*.

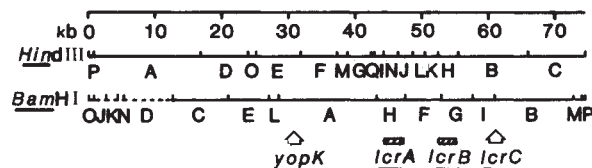


Fig. 1 Location of the mutation used in the present study within the physical map of pCD1. The site of the insertion mutation (*yopK*) is indicated for the *Bam*HI and *Hind*III restriction maps of pCD1. o shown are the locations of low- Ca^{2+} -response (*lcr*) genes necessary for the growth requirement for Ca^{2+} at 37 °C and for high expression of the V antigen. Broken lines indicate sections where the restriction fragment order is uncertain.

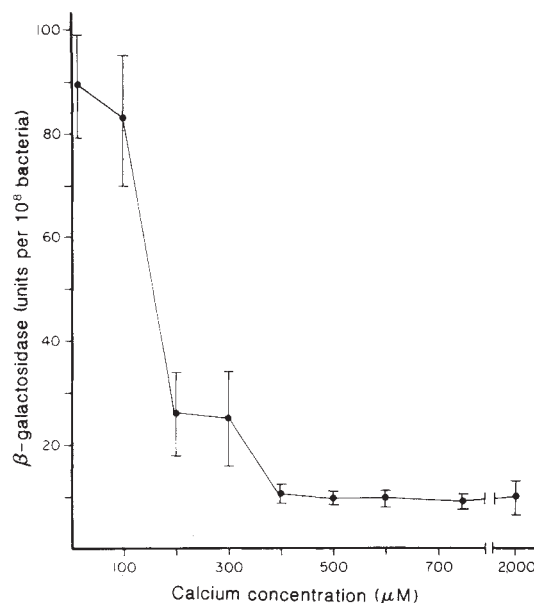


Fig. 2 Effect of Ca^{2+} on β -galactosidase production by *Y. pestis* KIM6 (pCD1::Mu d11-31). The enzyme activity is expressed as a function of the number of bacteria, as determined by counting in a Petroff-Hausser chamber under a microscope. One unit of β -galactosidase activity is equal to 10 μmol of *p*-nitrophenyl cleaved from the substrate *p*-nitrophenyl-D-galactopyranoside, per min, measured over a 12-h period. Each point represents the mean $\pm$ s.e.m. of 4-11 separate experiments, each run in duplicate.

Methods. *Y. pestis* KIM6 (pCD1::Mu d11-31), which had been maintained at 26 °C, was grown at 37 °C in heart infusion broth supplemented with 0.2% xylose, 2 mM MgCl_2 pH 7.0, at the indicated free Ca^{2+} concentration. The free Ca^{2+} concentration of solutions was calculated from the contaminating Ca^{2+} , measured with a Ca^{2+} -sensitive electrode (model F2112; Radiometer, Inc., Copenhagen, Denmark), and the amount of Ca^{2+} added. After 8 h of exponential growth, the bacteria were lysed with SDS/chloroform and β -galactosidase production was measured at pH 8.0 according to the method of Conchie *et al.*¹²

Y. pestis, the bacterium which causes plague, is a facultative intracellular organism which grows and multiplies in phagolysosomes within macrophages of the mammalian reticuloendothelial system¹. At least five virulence determinants have been identified for *Y. pestis*²; one of the most important of these is the response to low Ca^{2+} , which is manifested as two coordinately expressed properties: the requirement for millimolar Ca^{2+} for growth at 37 °C and the ability to synthesize a set of virulence-associated proteins, including the plague virulence antigens V and W (*Vwa*⁺ or *Lcr*⁺ phenotype)^{3,4}. While the precise functions of these proteins are unknown, mutational loss of the ability to express them results in loss of virulence. For example, the mouse

‡ To whom correspondence should be addressed.

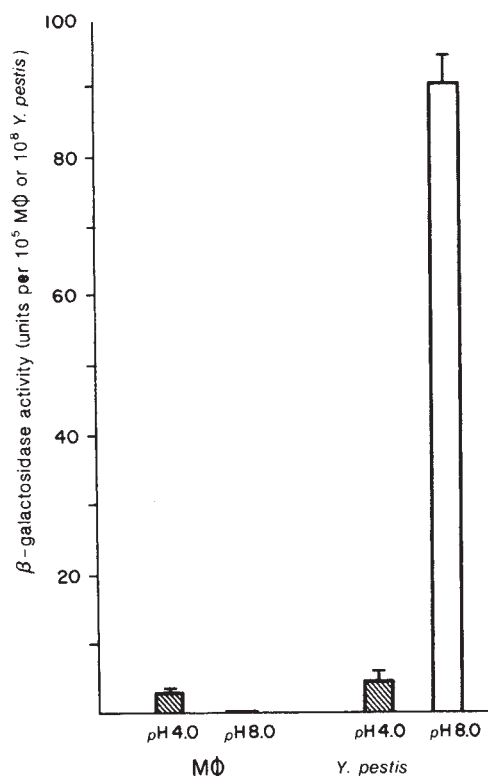


Fig. 3 Effect of pH on β -galactosidase activity of human macrophages and *Y. pestis* KIM6 (pCD1::Mu dI1-31).

Methods. *Y. pestis* KIM6 (pCD1::Mu dI1-31) was grown in the presence of $1.0 \mu\text{M}$ Ca^{2+} , as described in Fig. 2 legend. Human mononuclear cells were isolated from heparinized peripheral blood from healthy volunteers and cultured for 24 h in RPMI 1640 tissue culture medium supplemented with 10% fetal calf serum (FCS)¹³. All cells were lysed with SDS/chloroform, as described by Miller¹⁴, and β -galactosidase activity was measured as described by Conchie *et al.*¹² at pH 4.0 or pH 8.0. β -galactosidase activity is expressed in units (see Fig. 2).

LD₅₀ (50% lethal dose) for subcutaneously or intravenously injected Lcr⁺ *Y. pestis* is less than 10 bacteria, but $>10^7$ bacteria for Lcr⁻ *Y. pestis*. At 37 °C the expression of V and W antigens by *Y. pestis* is closely controlled by the calcium concentration in the growth environment, that is, the coordinated induction of these virulence proteins occurs only when the ambient calcium concentration is in the submillimolar range, well below the extracellular calcium concentration of the human host. Therefore, in order to express V and W antigens and remain virulent, *Y. pestis* must inhabit an intracellular environment with a low Ca^{2+} concentration. The unusual sensitivity of this organism to ambient calcium provides an opportunity to gain information about the phagolysosomal Ca^{2+} concentration. As precise measurement of V and W antigens is cumbersome, we have exploited an insertion mutant of the pCD1 plasmid which encodes the low- Ca^{2+} response, *Y. pestis* KIM6 [pCD1::Mu dI1(Ap lac)::Tn9-31], hereafter called *Y. pestis* KIM6 (pCD1::Mu dI1-31)⁵.

This mutant contains an insertion of the transposing operon fusion bacteriophage Mu dI1(Ap lac)⁶ that eliminates expression of YopK, an outer membrane protein necessary for virulence of *Y. pestis*⁵ (Fig. 1). The insertion does not affect the expression of V antigen or the growth requirement for Ca^{2+} at 37 °C (ref. 5). The insertion was stabilized against additional transposition or induction of phage functions by the presence of Tn9 in the Mu B gene⁵. Mu dI1(Ap lac) contains the structural genes of the *E. coli* lac operon, which are expressed as a result of readthrough from the yopK promoter. *Y. pestis* normally is lac⁻; accordingly, expression of the lacZ gene product, β -galac-

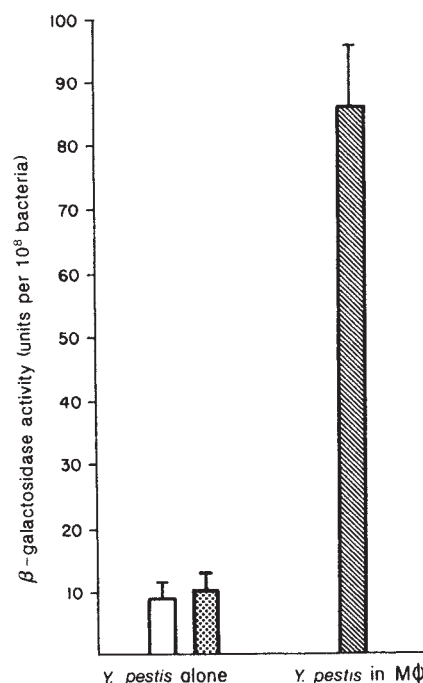


Fig. 4 β -Galactosidase production by extracellular *Y. pestis* recovered from dishes either without (open bar) or with (stippled bar) macrophages compared with β -galactosidase production by macrophage-associated *Y. pestis* (hatched bar). β -Galactosidase activity was assayed at pH 8.0 and is expressed as in Fig. 2. Results are the mean $\pm$ s.e.m. of five separate experiments each run in duplicate.

Methods. *Y. pestis* KIM6 (pCD1::Mu dI1-31), which had been grown at 37 °C in supplemented heart infusion broth containing 2.0 mM Ca^{2+} (ref. 1), were washed and resuspended in Hank's balanced salts solution (HBSS) containing 1.0 mM Ca^{2+} , 1.0 mM Mg^{2+} and 10% FCS, pH 7.4. The bacterial suspension (10^8 organisms) was added to Petri dishes in the absence or presence of a monolayer of human macrophages. The macrophage monolayers had been cultured at 37 °C in a humidified atmosphere of 95% air/5% CO_2 in RPMI-1640 medium plus 10% FCS for 24 h, and then washed with HBSS (as above) before addition of the bacterial suspension. In those dishes containing macrophages, the final ratio of bacteria to macrophages was $\sim 10:1$. After 8 h of incubation at 37 °C, extracellular *Y. pestis* were recovered by decanting the dishes, and the organisms were counted in a Petroff-Hausser chamber. The dishes with a macrophage monolayer were gently washed with prewarmed (37 °C) HBSS (1.0 mM Ca^{2+} , pH 7.4), leaving behind adherent macrophages containing intracellular and tightly attached *Y. pestis*. The cells were dislodged with a rubber policeman, counted in a haemocytometer and using an electronic particle counter (Coulter), and macrophage-associated *Y. pestis* were counted using the fluorescent acridine orange staining procedure of Bertalanffy *et al.*¹⁵.

tosidase, provides a measure of the transcription of yopK. Even though yopK is distant from the lcr genes (Fig. 1) and the V gene (S.C.S., unpublished), it shows the same pattern of expression as V antigen in yersiniae grown under different conditions⁵.

Figure 2 shows that the synthesis of β -galactosidase by *Y. pestis* KIM6 (pCD1::Mu dI1-31) is strictly controlled by the environmental calcium concentration. When grown at calcium concentrations between 400 and $2,000 \mu\text{M}$, a condition in which Lcr⁺ *Y. pestis* expresses V and W antigens only weakly, β -galactosidase production was suppressed. In growth media with a calcium concentration of $200\text{--}300 \mu\text{M}$, there was a small increase in β -galactosidase production. In contrast, ambient calcium concentrations between 1 and $100 \mu\text{M}$ induced maximal expression of β -galactosidase.

As our goal was to use bacterial β -galactosidase activity as an indicator of the phagolysosomal calcium concentration, it was important to distinguish between the β -galactosidase activity of the macrophage and that of *Y. pestis*; we therefore examined the pH optima for expression of macrophage and mutant *Y. pestis* β -galactosidase activity⁷ (Fig. 3). β -Galactosidase from disrupted macrophages was almost undetectable at pH 8.0, but fully active at pH 4.0. In contrast, β -galactosidase produced by the mutant *Y. pestis* grown in a low-calcium (1.0 μ M) environment was highly active at pH 8.0; at pH 4.0, less than 10% of the microbial β -galactosidase activity was detected.

Figure 4 compares β -galactosidase activity of extracellular mutant *Y. pestis* with that of organisms contained within human macrophages. In media containing 1 mM calcium, extracellular *Y. pestis* which had never been exposed to the macrophages produced little β -galactosidase. Similarly, extracellular *Y. pestis* recovered from the supernatant of dishes containing infected macrophages also produced little β -galactosidase. In contrast, the mutant *Y. pestis* contained within the human macrophages produced large quantities of β -galactosidase.

These studies, in which we used a genetically engineered intracellular pathogen to elucidate the environmental conditions in which it grows and multiplies *in vivo*, demonstrate that the intraphagolysosomal calcium concentration surrounding *Y. pestis* is sufficiently low to promote expression of calcium-sensitive virulence genes. We hypothesize that those genes encoding V antigen and outer membrane proteins are expressed during intraphagolysosomal growth of this pathogen; this was not expected, because it has been assumed that the *Y. pestis*-containing phagolysosome contains millimolar levels of free Ca^{2+} (ref. 1). Based on our data (Fig. 2), we predict that the intraphagolysosomal calcium concentration is <100 μ M. This is a surprising result in view of recent descriptions of ATP-dependent Ca^{2+} translocating pumps located in the plasma membrane of macrophages⁸ and neutrophils⁹ as well as a similar Ca^{2+} uptake pump in neutrophil lysosomes¹⁰. If the portion of the phagocyte plasma membrane which formed the phagocytic vacuole contained Ca^{2+} pumping activity and fused with lysosomes containing a similar Ca^{2+} pump, both of these ion translocating pumps would tend to raise the calcium concentration within this compartment. The consequences to intracellular *Y. pestis* would be the inability to express certain virulence genes and the organism would then be eradicated. It remains to be determined whether the calcium concentration is generally low in this intracellular compartment or whether *Y. pestis*, like *Toxoplasma gondii*¹¹, can modify its ambient growth environment to its own advantage.

We thank Robert Gelfand for technical assistance and Ms Debra Becky and Ms Carol McClarey for preparation of the manuscript. These studies were supported by NIH grants AI16732 and AI21017.

Diversity of murine gamma genes and expression in fetal and adult T lymphocytes

Joseph S. Heilig & Susumu Tonegawa

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The search for the genes encoding the T-cell receptor α and β chains revealed a third gene, $\text{T}\gamma$ (ref. 1), which shares with the $\text{T}\alpha$ (refs 2–7) and $\text{T}\beta$ (refs 8–15) genes a number of structural features, including somatic rearrangement during T-cell development. $\text{T}\gamma$ gene expression appears to be unnecessary in some mature T cells^{16,17} and is at its greatest in fetal thymocytes^{18,19}, encouraging speculation that $\text{T}\gamma$ has a role in T-cell development and may be involved in the recognition of polymorphic major histocompatibility complex (MHC) products during thymic education^{20,21}. One argument against the participation of $\text{T}\gamma$ in such a process has been its apparently limited diversity, due to the small number of gene segments available for rearrangement^{1,22}. We here describe the identification of additional $\text{T}\gamma$ V-gene segments and demonstrate that they can be rearranged to previously identified J- and C-gene segments and are expressed in fetal thymocytes. In addition we describe a variety of patterns of $\text{T}\gamma$ mRNA processing which may be significant for $\text{T}\gamma$ gene regulation.

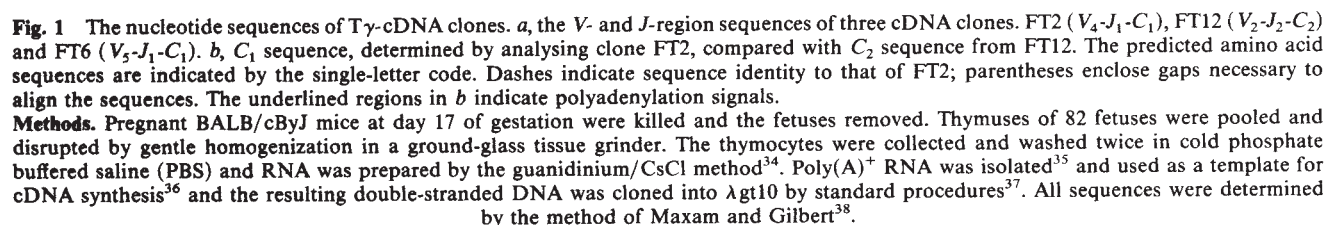
Previous studies^{1,22} identified three cross-hybridizing mouse V_γ -gene segments, V_1 , V_2 and V_3 , and three sets of cross-hybridizing J- and C-gene segments, $\text{J}_1\text{-C}_1$, $\text{J}_2\text{-C}_2$ and $\text{J}_3\text{-C}_3$ (see ref. 23 for nomenclature). Recently Iwamoto and co-workers²⁴ have reported a fourth set of J and C segments, $\text{J}_4\text{-C}_4$, which does not cross-hybridize with the other three (see Fig. 3). Although none of these germline gene segments have (with possible exception of C_3) obvious structural defects only the $\text{V}_2\text{-J}_2\text{-C}_2$ combination has been found as a complete rearranged gene^{1,16}. However, the existence of additional germline V_γ -gene segments, which do not cross-hybridize to the first three V-gene segments and are also rearranged in T cells, is suggested by two observations. First, many T-cell clones and splenic and peripheral T cells carry, in addition to a joined $\text{V}_2\text{-J}_2\text{-C}_2$ gene, an *EcoRI* DNA fragment containing a rearranged $\text{J}_1\text{-C}_1$ segment^{1,16,17,22}. Second, fetal thymocytes seem to contain more transcripts hybridizing to a C_γ -probe which detects C_1 , C_2 and C_3 sequences than to a V_γ -probe which detects V_1 , V_2 and V_3 sequences¹⁸.

To confirm this suspicion we constructed a cDNA library from fetal thymocytes and isolated cDNA clones which hybridize to a C_γ -probe but not to a V_γ -probe. Sequence analysis of one of these clones (FT2) revealed a new V_γ -gene segment (V_4) which shares 50% amino acid sequence homology with V_2 and is rearranged to the $\text{J}_1\text{-C}_1$ gene segment (Fig. 1). As is often the case with $\text{T}\gamma$ -gene transcripts^{1,16,23,25} the $\text{V}_4\text{-J}_1$ joint in FT2 results in a frame shift such that a translation product would terminate in the J region. Another clone, FT6, contains part of the sequence of another V_γ -gene segment (V_5) which is also rearranged to $\text{J}_1\text{-C}_1$ (Fig. 1). Although the complete sequence of V_5 was not contained in this clone, sufficient sequence was obtained to distinguish it from other V-gene segments. The $\text{V}_5\text{-J}_1$ joint in FT6 would also result in premature termination of translation. Nevertheless these new V_γ -gene segments indicate the potential of additional variability in the γ -gene family.

Figure 2 shows the result of a Southern blot analysis²⁶ in which DNA from a variety of T-cell clones and a 17-day-old fetal thymocyte population was analysed using C_2 , V_2 , V_4 and V_5 -probes in panels a, b, c and d, respectively. The C_2 -probe detects C_1 , C_2 and C_3 ; the V_2 -probe V_1 , V_2 and V_3 . The characteristic pattern of double rearrangement with C_2 and

Received 19 March; accepted 7 May 1986.

- Straley, S. C. & Harmon, P. A. *Infect. Immunity* **45**, 655–659 (1984).
- Brubaker, R. R. in *Microbiology—1979*, 168–171 (American Society of Microbiology, Washington, D.C., 1979).
- Brubaker, R. R. *Rev. Infect. Dis.* **5** (Suppl. 4), 748–758 (1983).
- Goguen, J. D., Yother, J. & Straley, S. C. *J. Bact.* **160**, 842–848 (1984).
- Straley, S. C. & Bowmer, W. S. *Infect. Immunity* (in the press).
- Casadaban, M. J. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4530–4533 (1979).
- Horst, J., Kluge, F., Beyreuther, K. & Gerok, W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3531–3535 (1975).
- Lew, P. D. & Stossel, T. P. *J. biol. Chem.* **255**, 5841–5846 (1980).
- Lagast, H., Lew, P. D. & Waldvogel, F. A. *J. clin. Invest.* **73**, 107–115 (1984).
- Klempner, M. S. *J. clin. Invest.* **76**, 303–310 (1985).
- Sibley, L. D., Weidner, E. & Krahenbuhl, J. L. *Nature* **315**, 416–419 (1985).
- Conchie, J., Findley, J. & Levvy, G. A. *Biochem. J.* **71**, 318–325 (1959).
- Klempner, M. S., Cendron, M. & Wyler, D. J. *J. infect. Dis.* **148**, 377–384 (1983).
- Miller, J. H. *Experiments in Molecular Genetics* 355 (Cold Spring Harbor Laboratory, New York, 1972).
- Bertalanffy, L. V., Masin, M. & Masin, F. *Cancer* **11**, 873–887 (1958).



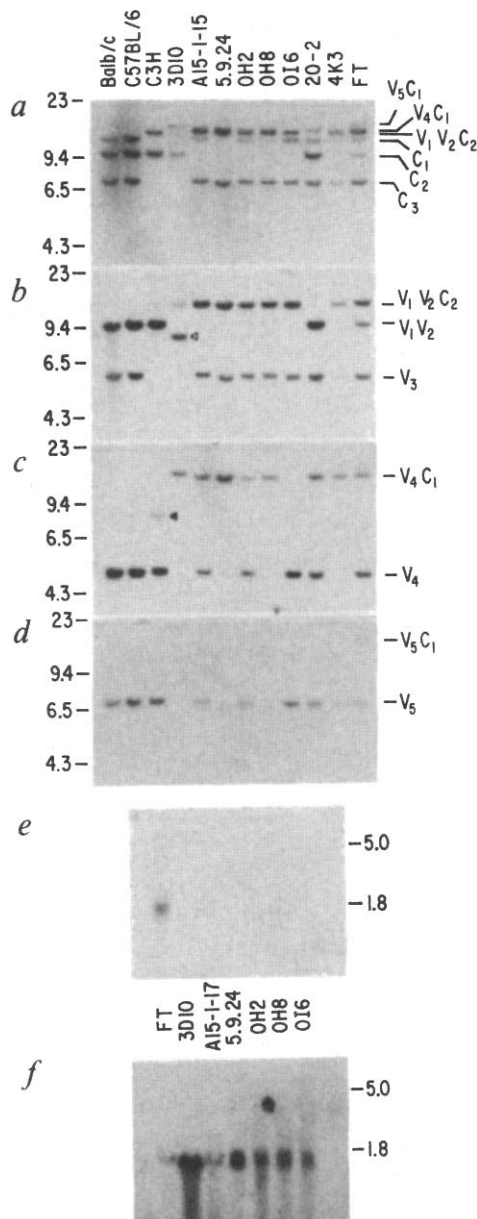


Fig. 2 Southern²⁶ and Northern³⁹ blot analyses of T-cell lines and fetal thymocytes. Panels *a-d* show Southern blot data while panels *e* and *f* show Northern blot data. In the Southern blots, BALB/c, C57BL/6 and C3H indicate kidney DNA of the designated strains; FT, DNA or RNA from 17-day-old BALB/c fetal thymocytes; 3D10, a KLH-specific suppressor T-cell line from C3H⁴⁰; A15-1-17, an alloreactive anti-I-A^k cytolytic T-cell line from A.TH⁴¹; 5-9-24, an I-A^d-ovalbumin-specific cytolytic/helper T-cell line from BALB/c (gift of B. Jones and C. Janeway); OH-2 and OH-8, D^b-H-Y-specific cytolytic T-cell lines from C57B/6 (gift of O. Kanagawa); OI-6, an I-A^b-H-Y-specific helper T-cell line from C57BL/6 (gift of O. Kanagawa); 20-2, an anti-H-2^b alloreactive helper T-cells line from BALB/c¹⁶; 4K3, an anti-L^d alloreactive cytolytic T-cell line from BALB.K²³. The unrearranged and rearranged gene segments are identified on the right hand side of panels *a-d*. The arrow in *b* indicates a V₂-cross-hybridizing band rearranged to unidentified sequence; probably an example of V₁-J₄-C₄ rearrangement (see Fig. 3 and ref. 24). The arrow in *c* indicates a V₄ cross-hybridizing band which may indicate partial digestion of the C3H kidney DNA.

Methods. In the Southern blot analysis, DNA samples were digested with *Eco*RI. Approximately 2 µg of each sample was electrophoresed in a 0.8% agarose gel, transferred to nitrocellulose and hybridized by standard procedures⁴². For the Northern blot analysis, RNA was prepared by the guanidinium/CsCl method³⁴ and approximately 12 µg of each sample was electrophoresed, transferred to nitrocellulose and hybridized as previously described^{17,39,42}. Probes used are: (*a*), the DNA encoding nearly the entire C₂ region depicted in Fig. 1*b* (detects C₁, C₂ and C₃); *b*, *f*, the region encoding the first 100 amino acids of V₂ (Fig. 1*a*) (detects V₁, V₂ and V₃); *c*, *e*, the region encoding the first 78 amino acids of V₄ (Fig. 1*a*); *d*, V₅ as described below. C₂, V₂ and V₄ probes were labelled by nick translation⁴³. Filters hybridized with these probes were washed in 0.4 × SSC/0.1% SDS at 65 °C. The V₅ probe was prepared and used as follows. A synthetic oligonucleotide corresponding to the 76 nucleotides of V₅ (Fig. 1*a*) was prepared and purified as described⁴⁴. A sesquidecamer complementary to the 3' end of this sequence (5' GCAGGCACAGTAGTA 3') was also synthesized. The probe was radiolabelled by a modification of the primer extension method⁴⁵. Hybridization was carried out at 55 °C. Filters were washed in 2 × SSC/0.1% SDS at 55 °C and exposed to film.

C₁ previously reported^{1,16} can be seen in most of these cells. In 20-2 however only C₁ rearrangement has occurred and in OI-6 only C₂ rearrangement. In cells containing a rearranged C₁ segment, the V₄ segment always rearranged to the same *Eco*RI fragment (Fig. 2*a, c*). The same correlation applies to C₂- and V₂-segments (Fig. 2*a, b*). By contrast, despite its rearrangement to the J₁-C₁ segment in the 17-day-old thymocytes, as evidenced by the Southern blots (Fig. 2*a, d*) and the isolation of the FT6 cDNA clone (Fig. 1), V₅ is not rearranged in any of the T-cell clones studied. Instead, V₅ is present only in those T-cell clones that retain at least one copy of unrearranged V₄- and C₁-segments, suggesting that rearrangement of V₄ to J₁-C₁ results in deletion of V₅. When these DNA samples are hybridized with a probe detecting a part of the C₄ gene, rearrangement is seen only in 3D10 DNA (data not shown, but apparently the same band as the one indicated by the arrow in Fig. 2*b*) which suggests that the V₁J₄C₄ rearrangement reported by Iwamoto and co-workers²⁴ is not very frequent among T-cell clones. The most probable organization of the various γ-gene segments, on the basis of all available data, is shown in Fig. 3.

The level of RNA containing V₂- or V₄-sequences was examined in 17-day-old fetal thymocytes as well as in a variety

of T-cell clones (Fig. 2*e, f*). All cells examined contained V₂ RNA. By contrast only fetal thymocytes contained V₄ RNA. Further analysis by cDNA cloning and sequencing revealed that most of the γ RNA in 17-day-old fetal thymocytes, whether of the V₂-J₂-C₂, V₄-J₁-C₁ or V₅-J₁-C₁ type, is apparently nonfunctional because of out-of-frame V-J junctions. Of eight cDNA clones sequenced which contain a rearranged V₂-, V₄- or V₅-segment, all had out-of-frame V-J junctions (see Fig. 1*a*, clones FT2, FT12 and FT6). Furthermore, many of these cDNA clones use splice sites other than those used in generating normal γ mRNA. Thus, as shown in Fig. 4, clone FT11 uses a splice acceptor in what is normally the V₂ coding region, clone FT6 employs a splice acceptor in what is normally the J₁-C₁ intron and clone FT5 uses the same acceptor and a splice donor also in the J₁-C₁ intron. In addition, clone FT10 presumably resulted from transcription initiating upstream of an unrearranged J₂-gene segment and clone FT13 uses an alternate polyadenylation site.

This study has added two V_γ- gene segments, V₄ and V₅, to the previously recognized sets of three V- and four J-C-gene segments. We now have evidence for rearrangement and expression at the RNA level of at least three types of Tγ genes,

Fig. 3 The arrangement of the γ genes, inferred from all available data. The orientation of the three gene clusters with respect to one another has not yet been determined. The broken line indicates the J_4 - C_4 cluster described elsewhere²⁴.

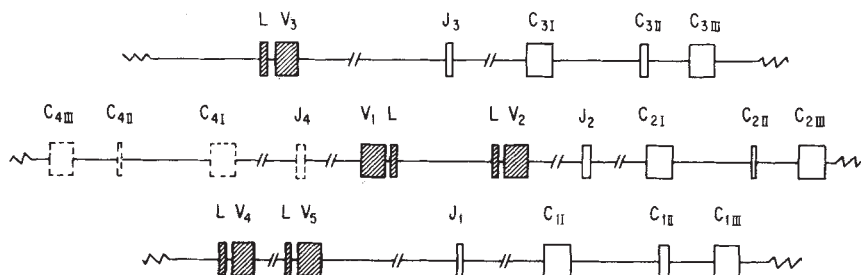
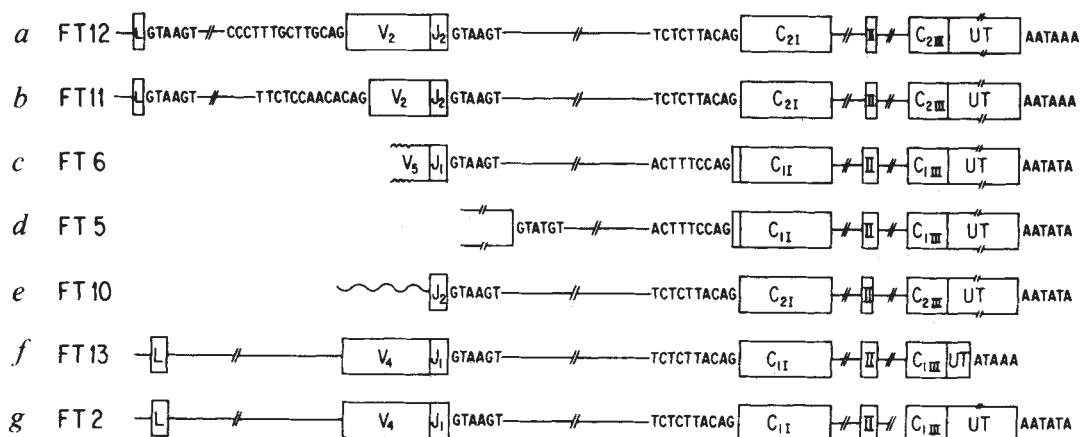


Fig. 4 Alternate splice patterns of γ -gene transcripts. Several types of mRNA generated by use of different splice signals are shown schematically. The boxed regions indicate exons and are labelled when the identity is known. The leader region is indicated by L; the first, second and third exons of the C regions by I, II and III respectively; the 3' untranslated regions by UT. Straight lines represent introns and are flanked by the nucleotide sequences of the splice signals. Only the sequences of relevant splice signals are indicated. In clone FT10 the wavy line indicates chromosomal sequence 5' of J_2 . The nucleotide sequence to the right of the UT region is the polyadenylation signal⁴⁶ presumed to be used by that clone. The normal splice patterns are depicted by clones FT12 and FT2. In clone FT6 the novel splice acceptor is 13 nucleotides upstream of the C_{1I} exon. In clone FT5 a splice donor in the J_1 - C_{1I} intron 248 nucleotides 5' of C_{1I} is used, as the acceptor described for clone FT6. The polyadenylation signals are underlined in Fig. 1b. We sequenced ten cDNA clones, two examples each of types a, b and g and one each of c, d, e and f. In each of the eight cDNA clones containing a V-J junction (FT5 and FT10 do not) the sequence at the V-J junction was different and resulted in frame shift. Splice signals were deduced by comparison to the consensus sequences suggested in ref. 47.



V_2 - J_2 - C_2 , V_4 - J_1 - C_1 and V_5 - J_1 - C_1 . Within each type there exists sequence variability in the V-J junction (this study and refs 16, 17, 23), indicating that the structural diversity of the T γ -gene products is not as limited as previously thought. Additional T γ -gene segments may exist which do not hybridize with the available γ -probes. In humans, at least two C_γ - and six V_γ -gene segments have been identified^{27,28}. The expanded T γ -gene diversity and the structural similarity of this gene to the T α and T β genes support the idea that the role of the T γ gene is in the recognition of polymorphic determinants on target cells. RNA transcripts detected by a C_2 or V_2 probe accumulate in the fetal thymocyte population, which is relatively rich in immature T cells, and are barely detectable in the resting T-cell population obtained from adult lymph nodes^{18,19}, suggesting that the V_2 - J_2 - C_2 gene product is involved in the early development of T cells, perhaps in the interaction with the polymorphic MHC determinants presented to immature T cells by thymic epithelial cells^{29,30}. The present finding that V_4 - J_1 - C_1 and V_5 - J_1 - C_1 transcripts are present in fetal thymocytes but not in a variety of T-cell clones supports this hypothesis. However, the fourth T γ gene, V_1 - J_4 - C_4 , does not appear to be used abundantly in 17 day-old fetal thymocytes preferentially because it is not represented in our cDNA library, as shown by the fact that all clones hybridizing to a V_2 - probe also hybridize to a C_1 probe (V_2 and V_1 cross hybridize).

The present study shows that most of the T γ RNA present in 17-day-old fetal thymocytes is defective due to out-of-frame V-J joining and abnormal splicing. Examples of out-of-frame joining and abnormal splicing of T-cell-receptor genes have been reported^{4,17,25}. However, the strong bias in favour of presumably nonfunctional mRNA for T γ is striking. Since it is difficult to imagine that in-frame V-J joining is specifically suppressed, the high incidence of out-of-frame cDNA may reflect an aspect of regulation of T γ -gene expression. It is

possible that the γ protein is required only during an early phase of T-cell development and its continued presence is detrimental to the cell. If so, T cells will be more tolerant to continued transcription of an out-of-frame T γ gene than an in-frame gene. The implication of this view is that the 17-day-old fetal thymocyte population is dominated by immature T cells that have already passed through the stage in which the T γ gene plays a critical role. Alternatively, the functional expression of the T γ gene is restricted to a small fraction of thymocytes destined to become a subset of T-cells as yet unidentified. The significance of the high incidence of apparently abnormal RNA splicing is unclear but may also reflect the need to eliminate in-frame T γ gene products once the cell has passed a critical stage in development. Precedents exist in the literature for the role of alternative splicing in the inactivation of a gene function during viral and cellular development³¹⁻³³.

We thank R. Garman, D. Raulet and Y. Takagaki for helpful advice, J. McMaster, H. Siegel and D. Wilson for technical assistance, H. N. Eisen, M. Iwashima, C. Janeway, B. Jones, O. Kanagawa, D. Kranz, M. Pierres, E. Reilly and T. Tada for the T cell lines used in this study and W. A. Gilbert for assistance in preparation of Fig. 1, N. Hopkins for comments on the manuscript and Elly Basel for preparing the manuscript. This work was supported by NIH grants CA28900 and AI17879 to S.T. and P30-CA14051-14 (a core grant to P. Sharp).

Note added in proof: After submission of this paper other groups have reported the identification and sequencing of additional V_γ -gene segments: V_4 in refs 48, 49; V_4 , V_5 , V_6 in ref. 50

Received 30 April; accepted 6 June 1986.

1. Saito, H. *et al. Nature* **309**, 757-762 (1984).
2. Saito, H. *et al. Nature* **312**, 36-39 (1984).
3. Chien, Y.-h. *et al. Nature* **312**, 31-35 (1984).
4. Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
5. Hayday, A. C. *et al. Nature* **316**, 828-832 (1985).

6. Winoto, A., Mjolsness, S. & Hood, L. *Nature* **316**, 832-836 (1985).
7. Yoshikai, Y. *et al.* *Nature* **316**, 837-840 (1985).
8. Yanagi, Y. *et al.* *Nature* **308**, 145-149 (1984).
9. Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149-153 (1984).
10. Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
11. Chien, Y.-h., Gascoigne, N., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
12. Siu, G. *et al.* *Cell* **37**, 393-401 (1984).
13. Mallissen, M. *et al.* *Cell* **37**, 1101-1110 (1984).
14. Gascoigne, N., Chien, Y.-h., Becker, D., Kavalier, J. & Davis, M. M. *Nature* **310**, 387-391 (1984).
15. Siu, G. *et al.* *Nature* **311**, 344-350 (1984).
16. Kranz, D. M. *et al.* *Nature* **313**, 752-755 (1985).
17. Heilig, J. S. *et al.* *Nature* **317**, 68-70 (1985).
18. Raulet, D. H., Garman, R. D., Saito, H. & Tonegawa, S. *Nature* **314**, 103-107 (1985).
19. Snodgrass, H. R., Dembic, Z., Steinmetz, M. & vonBoehmer, H. *Nature* **315**, 232-233 (1985).
20. Zinkernagel, R. M. *Immun. Rev.* **42**, 224-270 (1978).
21. Bevan, M. J. & Fink, P. J. *Immun. Rev.* **42**, 3-19 (1978).
22. Hayday, A. C. *et al.* *Cell* **40**, 259-269 (1985).
23. Reilly, E. B., Kranz, D. M., Tonegawa, S. & Eisen, H. N. *Nature* **321**, 878-880 (1986).
24. Iwamoto, A. *et al.* *J. exp. Med.* **163**, 1203-1212 (1986).
25. Dyalnas, D. P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 2619-2623 (1986).
26. Southern, E. J. *molec. Biol.* **98**, 503-517 (1975).
27. Lefranc, M.-P. & Rabbitts, T. H. *Nature* **316**, 464-466 (1985).
28. Quettermous, T. *et al.* *Science* **231**, 252-255 (1986).
29. Jenkinson, E. J., Franchi, L., Kingston, R. & Owen, J. J. T. *Eur. J. Immun.* **12**, 583-587 (1982).
30. Lo, D. & Sprent, J. *Nature* **319**, 672-675 (1986).
31. Sharp, P. A. in *The Adenoviruses* (ed. Ginsberg, H. S.) 173-204 (Plenum, New York, 1984).
32. Laski, F. A., Rio, D. C. & Rubin, G. M. *Cell* **44**, 7-19 (1986).
33. Rio, D. C., Laski, F. A. & Rubin, G. M. *Cell* **44**, 21-32 (1986).
34. Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
35. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
36. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
37. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A Practical Approach* (ed. Glover, D.) 49-78 (IRL, Oxford, 1984).
38. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
39. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
40. Nakauchi, H., Ohno, I., Kim, M., Okumura, K. & Tada, T. *J. Immun.* **132**, 88-94 (1984).
41. Pierres, A., Schmitt-Verhulst, A. M., Bufe, M., Golstein, P. & Pierres, M. *Scand. J. Immun.* **15**, 619-625 (1982).
42. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York).
43. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. mol. Biol.* **113**, 237-256 (1977).
44. Reilly, E. B., Reilly, R. M., Breyer, R. M., Saver, R. T. & Eisen, H. N. *J. Immun.* **133**, 471-475 (1984).
45. Studencki, A. B. & Wallace, R. B. *DNA* **3**, 7-15 (1984).
46. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).
47. Breathnach, R. & Chambon, P. A. *Rev. Biochem.* **50**, 349-383 (1981).
48. Trautner, A. *et al.* *EMBO J.* **5**, 1584-1594 (1986).
49. Rupp, F., Frech, G., Hengartner, H., Zinkernagel, R. M. & Joho, R. *Nature* **321**, 876-878 (1986).
50. Garman, R., Doherty, P. J. & Raulet, D. H. *Cell* **45**, 733-742 (1986).

A novel V_H to V_HDJ_H joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production

Michael Reth, Peter Gehrman, Eva Petrac & Petra Wiese

Institute for Genetics, University of Cologne, Weyertal 121, D-5000 Köln 41, FRG

During B-cell development, the V_H genes of immunoglobulin heavy (H) chains are assembled from three different germline components: the variable (V_H) segment, the diversity (D) segment and the joining (J_H) segment^{1,2}. The joining between two segments involves the recognition of conserved nonamer-heptamer sequences bordering each segment, double-stranded cuts at the heptamer-segment border, and the re-ligation of the two segment ends which have frequently been modified by the deletion and addition of nucleotides³⁻⁶. The flexibility of the joint increases V_HDJ_H variability. However, it also results in many pre-B cells which do not produce immunoglobulin H chains and have non-functional V_HDJ_H complexes carrying the V_H and J_H coding sequences in different reading frames⁷. We show here that such 'null cells' are not dead-end products of the B-cell developmental pathway but can perform a novel V_H to V_HDJ_H joining using a 5' V_H segment to replace the V_H sequence of the V_HDJ_H complex. This process can result in the generation of a $V_HDJ_H^+$ complex and the subsequent expression of an immunoglobulin heavy chain.

In the past we have studied the order and control of immunoglobulin gene rearrangements in the Abelson pre-B-cell line 300-19 (refs 8-11) derived from the bone marrow of an outbred NIH/Swiss mouse¹². 300-19 cells which originally carry a DJ_H3 complex on each J_H allele⁸ undergo D to J_H and V_H to DJ_H assembly while growing in culture⁹. Approximately 30% of the V_H to DJ_H joints of 300-19 place the V_H and J_H coding sequence in the same reading frame thus resulting in the production of a μ -chain⁹, the first H chain class expressed in pre-B cells. Many of the 300-19 subclones, however, contain $V_HDJ_H^-$ complexes and fail to produce a μ -chain. To study the fate of such null cells, we analysed P17 and Pa2-5, two isolates of 300-19 carrying unproductive V_HDJ_H complexes.

Aberrant V_H to DJ_H joints had obviously occurred on both J_H alleles of P17 because we could not detect any μ -chains nor 5' D sequences in P17 (data not shown). Indeed, all 5' D sequences are generally deleted during the V_H to DJ_H joint of a J_H allele³. To our surprise, P17 became more and more μ -

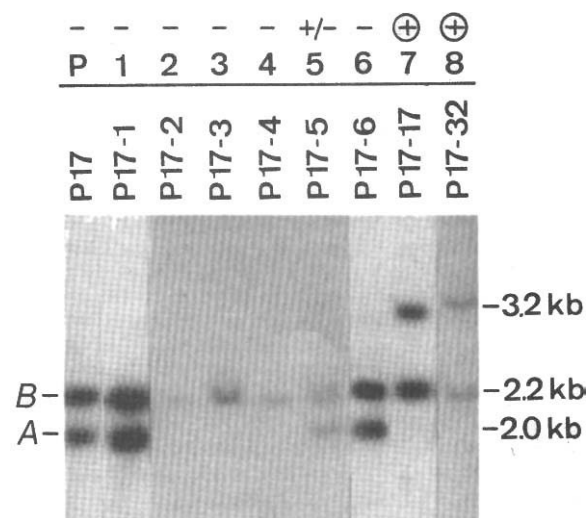


Fig. 1 J_H rearrangement in subclones of P17. The two parental V_HDJ_H alleles of P17 are labelled A and B. Approximately 10 μ g of genomic DNA of the different subclones was digested with *EcoRI*, electrophoresed through 1% agarose, blotted onto nitrocellulose and assayed for hybridization to a ³²P-labelled J_H probe as described previously^{3,8}. μ Production in the subclones was determined by Western blotting (see Fig. 2) and is indicated by + or -.

positive over a 3-4 week culture period. After 4 weeks of culture we subcloned P17 by limiting dilution and analysed μ production and J_H rearrangement of isolated subclones. Of 55 P17 subclones analysed, 8 (15%) were found strongly μ -positive in a Western dot assay (data not shown). On a Southern blot the two V_HDJ_H alleles of P17 (A and B) are visible as J_H -positive *EcoRI* fragments of 2 and 2.2 kilobases (kb) respectively (Fig. 1, lane P). In the μ -positive subclones P17-17 and P17-32, the 2-kb *EcoRI* fragment of the P17-A allele had undergone rearrangement and appeared in a new J_H -positive *EcoRI* fragment of 3.2 kb (Fig. 1, lanes 7, 8). Three of the six randomly picked subclones had deleted the 2-kb *EcoRI* fragment (Fig. 1, lanes 2-4). Another of these subclones (P17-5) with weak μ expression (data not shown) showed a faint band of 3.2 kb (Fig. 1, lane 5), suggesting that part of the P17-5 culture was undergoing the same rearrangement event as that occurring in P17-17 and P17-32. The 2.2-kb *EcoRI* fragment of the P17-B allele stayed unchanged in all P17 subclones. Thus only the V_HDJ_H complex of the A allele of P17 had undergone genetic alterations. These were either its complete deletion or a further rearrangement correlated with μ production.

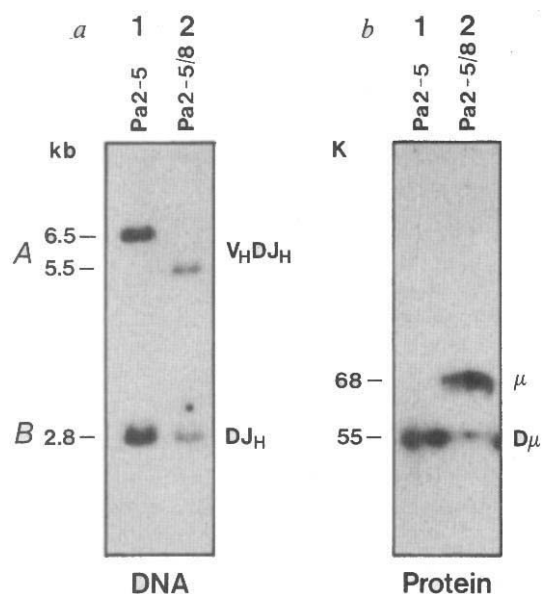
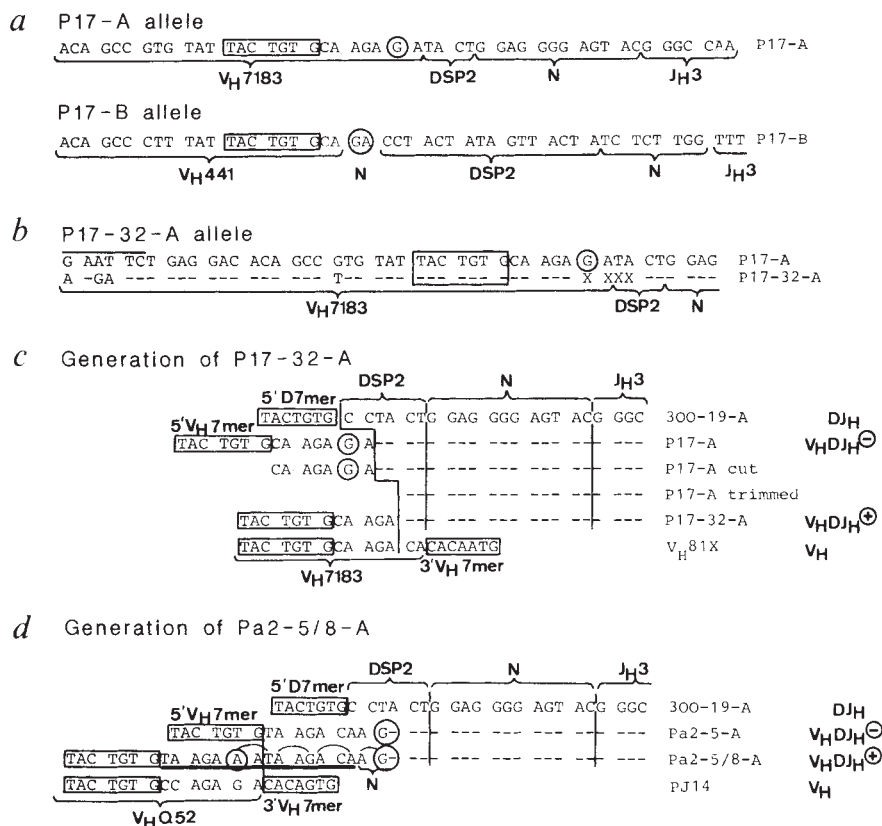


Fig. 2 J_H rearrangement and μ production in subclones Pa2-5 and Pa2-5/8. *a*, Approximately 10 μ g of genomic DNA was double digested with *Bam*HI and *Eco*RI and probed for hybridization with a J_H probe as described previously^{3,8}. The position of the two J_H alleles is indicated by A and B. *b*, Intracellular protein of 10^7 cells of each subclone was fractionated by electrophoresis through 10% polyacrylamide gels and electrotransferred to nitrocellulose. Bound μ protein was detected by 125 I-labelled antibodies as described previously⁸.

Fig. 3 Nucleotide sequence of V_H to DJ_H and V_H to V_HDJ_H rearrangements in 300-19 subclones. The heptamer recognition sequence is boxed. Sequences derived from V_H , D , N and J_H3 are labelled accordingly. The V_H sequences are given the names of their respective V_H family as described by Brodeur and Riblet^{13,14}. Circled nucleotides mark the position where the V_H and DJ_H reading frames are not matched (out of phase). *a*, Sequence of the two V_HDJ_H alleles (A and B) of P17. *b*, Comparison of the P17-32-A sequence with that of P17-A. The P17-A sequence is given and the identity of the P17-32-A sequence is indicated by a dash. Nucleotide deletions in the P17-32-A sequence are indicated by crosses. An internal *Eco*RI site in the V_H sequence of P17-A is overlined, indicating the position up to which both sequences could be compared. *c*, Generation of P17-32. The sequence of the 300-19 DJ_H3 A allele⁸ is given in the first line and the remaining $D-N-J_H3$ sequence in the V_HDJ_H3 complexes is indicated by dashes. All V_H sequences are shown up to the border of the internal V_H heptamer. The V_H sequence of P17-32-A is compared with that of V_H81X (ref. 41) a germline V_H element of the V_H7183 family. *d*, Generation of Pa2-5/8. $V_H-D-N-J_H3$ sequences are presented as described above. The partial duplication of the seven terminal V_H nucleotides is underlined. The reading frame of the V_H and J_H coding sequence at the V_H to V_HDJ_H joint is indicated above the sequence. The V_H sequence of Pa2-5/8 is compared with that of PJ14 (ref. 4), a germline V_H segment of the V_HQ52 family.



Rearrangement of a V_HDJ_H complex was also seen in Pa2-5, another isolate of the 300-19 culture. Pa2-5 carried an aberrant V_H to DJ_H rearrangement on the A allele and a parental (300-19) DJ_H3 complex on the B allele. The DJ_H3 -B allele is expressed in Pa2-5 as well as in 300-19 (ref. 8) and gives rise to a short $D\mu$ -chain of relative molecular mass, 55,000 (55K; Fig. 2b, lane 1). After 2 weeks of culture, we subcloned Pa2-5 and randomly chose 12 subclones to analyse their J_H rearrangements and μ production. One of these subclones, Pa2-5/8, expresses both μ and $D\mu$ -chains (Fig. 2b, lane 2). The presence of the $D\mu$ -chain suggested that the μ expression of Pa2-5 did not result from a regular V_H to DJ_H3 rearrangement on the B allele but rather from an alteration of the V_HDJ_H A allele. A Southern blot analysis confirmed this interpretation (Fig. 2a). The 2.8-kb *Bam*HI/*Eco*RI fragment of the DJ_H B allele was present in both lines while the 6.5-kb fragment of the Pa2-5 A allele was rearranged to a J_H -positive fragment of 5.5 kb in Pa2-5/8 (Fig. 2a, lanes 1, 2). Reprobing the Southern blot with V_H -specific probes demonstrated that the V_HDJ_H complexes of both Pa2-5 and Pa2-5/8 contained a V_HQ52 element (data not shown). Because we have characterized and mapped most V_HQ52 elements used by 300-19 cells⁹, we could identify the rearranged V_H elements of Pa2-5 and Pa2-5/8 as being the two most 3' situated V_H segments of the V_HQ52 family (for definition of V_H families see refs 13-15). This suggested that the V_HQ52 sequence in the V_HDJ_H complex of Pa2-5 was replaced in Pa2-5/8 by a neighbouring V_H segment.

To further characterize the unexpected rearrangement in the P17 and Pa2-5 culture, we cloned the V_HDJ_H alleles of P17 and Pa2-5 as well as the newly generated productive rearrangement of P17-32 and Pa2-5/8. The sequence of the P17 J_H alleles (Fig. 3a) confirmed that both had assembled a V_HDJ_H complex by joining either a V_H7182 (allele A) or a V_H441 element (allele B) to the previously sequenced DJ_H3 complexes of 300-19⁸.

Methods. All V_HDJ_H alleles were cloned as J_H -positive *Eco*RI fragments in either CH16A or CH35⁴². Restriction fragments carrying the respective rearrangement were subcloned into pUC-19 and sequenced according to Maxam and Gilbert⁴³. The V_H family of the rearranged V_H segments was determined by a comparison of their amino-acid sequence to a representative collection of V_H sequences¹⁵ and by a blotting analysis using V_H -specific probes⁴¹.

Both $V_H DJ_H$ alleles of P17 carry the V_H and J_H coding sequences in different reading frames (see circled nucleotides in Fig. 3a), thus explaining the absence of μ production in P17. A similar out of phase joint has also occurred in the V_H to DJ_H rearrangement of Pa2-5 (Fig. 3d). The rearranged A allele of P17-32 carries a $V_H DJ_H^+$ complex which obviously was derived from the $V_H DJ_H$ complex of P17-A because both complexes contain identical $D-N-J_H3$ sequences (Fig. 3b, c). The $V_H DJ_H^+$ complex of P17-32-A also contains a V_H7183 segment whose determined sequence (see legend to Fig. 3b), however, differs from that of P17-A by an exchange of four nucleotides and a deletion of four nucleotides at the V_H-D border (Fig. 3b).

How can the new $V_H DJ_H$ complex of P17-32-A have been generated from that of P17-A? Although alternative explanations exist¹⁶, we suggest that the P17-32-A $V_H DJ_H^+$ complex results from a V_H to $V_H DJ_H$ joining event. The V_H sequence of P17-A as well as that of most V_H elements contains the conserved heptamer TACTGTG seven nucleotides away from the 3' end (boxed in Fig. 3). As also noted by others¹⁷, the sequence of this internal V_H heptamer is identical to that of 5' D heptamers (see Fig. 3c). Therefore, like a DJ_H complex, a $V_H DJ_H$ complex carries a 5' heptamer and thus could be partner in the joining to a new V_H segment. Figure 3c depicts the proposed steps in such a V_H to $V_H DJ_H$ joining to generate P17-32-A. The joining would start with a precise cut in P17-A at the very end of the internal V_H heptamer, followed by the terminal deletion of seven V_H and two D nucleotides and the addition of a new 5'-located V_H7183 element to the remaining $D-N-J_H3$ sequences. In Pa2-5/8 the V_H to $V_H DJ_H$ joint obviously occurred without any nucleotide deletion or addition (Fig. 3d, see vertical line between the heptamers), resulting in a $V_H DJ_H^+$ complex with a partial duplication of the last seven V_H nucleotides (underlined in Fig. 3d). Both the V_H to $V_H DJ_H$ joints described brought the V_H and J_H coding sequence back into the same reading frame and thus allowed μ expression.

The consensus recognition sequence consists of a conserved heptamer and nonamer separated by a spacer of either 12 or 23 base pairs and joining normally occurs only between segments flanked by recognition sequences of different spacer length (12/23 joining rule)^{1,4}. The V_H to $V_H DJ_H$ joining seems to violate this rule because it employs an isolated heptamer in the V_H sequence of the $V_H DJ_H$ complex. Joining events involving an isolated heptamer, however, also occur at the mouse and human κ locus and seem to be a possible variation of the rearranging process¹⁸⁻²⁰. The V_H replacement of Pa2-5/8 seems to involve two neighbouring V_H segments and the same may be true for P17-32 where the V_H to $V_H DJ_H$ joint did not result in the deletion of many V_H7183 germline segments (data not shown). The use of neighbouring V_H segments in the V_H to $V_H DJ_H$ joint may be connected with the local activation of V_H germline transcripts on a $V_H DJ_H$ allele, and stresses the known correlation between

transcription and rearrangement at the immunoglobulin loci²¹⁻²⁶. Wang and Calame²⁷ have described the transcription of a normally silent V_H germline segment in B cells which carry a $V_H DJ_H$ complex 16 kb downstream from this V_H segment. The upstream V_H transcription is thought to be activated once the V_H to DJ_H joint has placed the immunoglobulin H gene enhancer element close to the 5'-located V_H segment. The local action of the H gene enhancer²⁸⁻³⁰ may also be an explanation for the local activation of the V_H to $V_H DJ_H$ joint. Thus our studies support the idea of a dual function of the H gene enhancer, namely to provide transcriptional as well as recombinational enhancement^{2,31}.

By completely replacing one V_H sequence by another (see P17-32), a V_H to $V_H DJ_H$ joint may generate a $V_H DJ_H$ sequence indistinguishable from that derived from a V_H to DJ_H joint. Thus the importance and frequency of V_H to $V_H DJ_H$ joints during B-cell development cannot easily be evaluated from an analysis of published $V_H DJ_H$ sequences³². In the two subclones analysed, however, V_H to $V_H DJ_H$ joining seems to occur with a frequency similar to that of V_H to DJ_H rearrangements in the 300-19 culture⁹. Furthermore, the internal V_H heptamer which mediates the V_H to $V_H DJ_H$ joint is highly conserved in most mouse and human V_H segments³². Specifically, the heptamer is found in V_H segments of all mouse V_H families except those of V_H J606, the most 5' located V_H family¹⁴. Most of the V segments of other gene loci³³⁻³⁵ (those of the immunoglobulin light chains or T-cell receptor chains) do not carry an internal V heptamer with the interesting exception of the T-cell receptor γ -chain where the heptamer has been found in all mouse and human V_γ segments sequenced³⁶⁻³⁸.

The strong conservation of the internal V_H heptamer suggests that V_H to $V_H DJ_H$ joining is important for the generation of functional $V_H DJ_H$ complexes at the V_H locus. Indeed, a V_H to DJ_H joint can occur only once on a J_H allele because all germline D segments which are mediating the V_H to J_H joint are deleted during the joining process³. Statistically, two out of three V_H to DJ_H joints place the V_H and J_H coding sequences in different reading frames, generating up to 50% of null pre-B cells with aberrant $V_H DJ_H$ complexes on both J_H alleles. The V_H to $V_H DJ_H$ joining mechanism may have specifically evolved to rescue these null pre-B cells. V_H replacements could occur repeatedly on a J_H allele, increasing its chance of generating a productive $V_H DJ_H$ complex which leads to the expression of a μ -chain. V segments of other V-gene loci (see above) may not need a V replacement mechanism because they have the possibility of multiple V to J joints^{39,40}. The consequences of the V_H to $V_H DJ_H$ joining mechanism for allelic exclusion, V_H usage and generation of diversity in developing B cells are being investigated.

We thank Lise Leclercq and Klaus Rajewsky for critical reading of this manuscript. This work was supported by BMFT grant BCI-0365/2, Project 12.

Received 31 March; accepted 21 May 1986.

1. Tonegawa, S. *Nature* **302**, 575-581 (1983).
2. Yancopoulos, G. D. & Alt, F. W. *A. Rev. Immunol.* **4**, 339-368 (1986).
3. Alt, F. W. *et al. EMBO J.* **3**, 1209-1219 (1984).
4. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
5. Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
6. Steinmetz, M., Altenburger, W. & Zachau, H. G. *Nucleic Acids Res.* **8**, 709-720 (1980).
7. Hagiya, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 145-149 (1986).
8. Reth, M. G. & Alt, F. W. *Nature* **312**, 418-423 (1984).
9. Reth, M. G., Jackson, S. & Alt, F. W. *EMBO J.* (in the press).
10. Reth, M. G., Ammirati, P., Jackson, S. & Alt, F. W. *Nature* **317**, 353-355 (1985).
11. Alt, F. W., Blackwell, T. K., DePinho, R. A., Reth, M. G. & Yancopoulos, G. D. *Immun. Rev.* **89**, 5-20 (1986).
12. Rosenberg, N. & Baltimore, D. *J. exp. Med.* **143**, 1453-1463 (1976).
13. Brodeur, P. & Riblet, R. *Eur. J. Immunol.* **14**, 922-930 (1984).
14. Brodeur, P. & Riblet, R. *UCLA Symp. molec. cell. Biol. new Ser.* **18**, 445 (1984).
15. Dildrop, R. *Immun. Today* **5**, 85-86 (1984).
16. Krawinkel, U., Zechlein, G., Brüggemann, M., Radbruch, A. & Rajewsky, K. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4997-5001 (1983).
17. Early, P., Nottenburg, C., Weissman, I. & Hood, L. *Molec. cell. Biol.* **2**, 829-831 (1982).
18. Hochli, J. & Zachau, H. G. *Nature* **302**, 260-263 (1983).
19. Moore, M. W., Durdik, J., Persiani, D. M. & Selsing, E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6211-6215 (1985).
20. Simionovitch, K. A., Bakshi, A., Goldman, P. & Korsmeyer, S. J. *Nature* **316**, 260-263 (1985).
21. Kemp, D. J., Harris, A. W., Cory, S. & Adams, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2876-2880 (1980).
22. Van Ness, B. G. *et al. Cell* **27**, 593-602 (1981).
23. Picard, D. & Schaffner, W. *EMBO J.* **3**, 3031-3039 (1984).
24. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271-281 (1985).
25. Blackwell, T. K., Yancopoulos, G. D. & Alt, F. W. *UCLA Symp. molec. cell. Biol., new Ser.* **19**, 537 (1984).
26. Siu, G. *et al. Nature* **311**, 344-350 (1984).
27. Wang, X. F. & Calame, K. *Cell* **43**, 659-665 (1985).
28. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-728 (1983).
29. Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
30. Neuberger, M. S. *EMBO J.* **2**, 1373-1378 (1983).
31. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L. & Alt, F. W. *Cell* **44**, 251-259 (1986).
32. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. (eds) *Sequences of Proteins of Biological Interest* No. 80-2008 (NIH, Bethesda, 1983).
33. Barth, R. K. *et al. Nature* **316**, 518-523 (1985).
34. Arden, B., Klotz, J., Siu, G. & Hood, L. *Nature* **316**, 518-523 (1985).
35. Becker, D. *et al. Nature* **317**, 430-434 (1985).
36. Kranz, D. M. *et al. Nature* **313**, 752-755 (1985).
37. Hayday, A. C. *et al. Cell* **40**, 259-269 (1985).
38. Lefranc, M. P., Forster, A. & Rabbitts, T. H. *Nature* **319**, 420-422 (1986).
39. Lewis, S., Rosenberg, N., Alt, F. & Baltimore, D. *Cell* **30**, 807-816 (1982).
40. Feddersen, R. M. & Van Ness, B. G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4793-4797 (1985).
41. Yancopoulos, G. D. *et al. Nature* **311**, 727-733 (1984).
42. De Wet, T. *et al. J. Virol.* **33**, 401-410 (1980).
43. Maxam, A. & Gilbert, W. *Methods Enzymol.* **65**, 419-460 (1980).

Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1⁺ B-cell lymphoma

Robert Kleinfield*, Richard R. Hardy†, David Tarlinton†, Jeffery Dangle†, Leonard A. Herzenberg† & Martin Weigert*

* Institute for Cancer Research, Philadelphia, Pennsylvania 19111, USA

† Department of Genetics, Stanford University, Stanford, California 94305, USA

The early stages of murine B-cell differentiation are characterized by a series of immunoglobulin gene rearrangements which are required for the assembly of heavy(H)- and light(L)-chain variable regions from germline gene segments. Rearrangement at the heavy-chain locus is initiated first and consists of the joining of a diversity (D_H) gene segment to a joining (J_H) gene segment. This forms a DJ_H intermediate to which a variable (V_H) gene segment is subsequently added. Light-chain gene rearrangement follows and consists of the joining of a V_L gene segment to a J_L gene segment: once a productive light-chain gene has been formed the cell initiates synthesis of surface immunoglobulin M (sIgM) receptors (reviewed in ref. 1). These receptors are clonally distributed and may undergo further diversification either by somatic mutation^{2,3} or possibly by continued recombinational events⁴. Such recombinational events have been detected in the Ly 1⁺ B-cell lymphoma NFS-5, which has been shown to rearrange both λ and H-chain genes subsequent to the formation of sIgM ($\mu\kappa$) molecules⁵. Here we have analysed a rearrangement of the productive allele of NFS-5 and found that it is due to a novel recombination event between V_H genes which results in the replacement of most or all of the coding sequence of the initial V_HQ52 rearrangement by a germline V_H7183 gene. Embedded in the V_H coding sequence close to the site of the cross-over is the sequence 5' TACTGTG 3', which is identical to the signal heptamer found 5' of many D_H gene segments⁶. This embedded heptamer is conserved in over 70% of known V_H genes⁷⁻¹⁷. We suggest that this heptamer mediates V_H gene replacement and may play an important part in the development of the antibody repertoire.

The Ly 1⁺ B-cell lymphoma NFS-5 was generated by inoculating a newborn NFS/N mouse with ecotropic murine leukaemia virus (Cas 2SM)¹⁸. Recent studies have indicated that this lymphoma is capable of expressing a number of distinct surface immunoglobulin phenotypes during growth *in vitro*⁵. Analysis of primary cultures of NFS-5 on the fluorescence-activated cell sorter (FACS) shows a $\mu^- \kappa^-$ population from which $\mu^+ \kappa^-$ cells spontaneously arise. Kappa light-chain synthesis can be induced by treatment with the B-cell mitogen lipopolysaccharide (LPS), to give a $\mu^+ \kappa^+$ population. Continued growth of these κ -expressing cells in LPS results in the appearance of a small population, comprising 1-5% of the culture, expressing λ light chains. These λ -expressing cells generally have a $\mu^+ \lambda^+$ surface phenotype, however $\mu^+ \kappa^+ \lambda^+$ co-expressors have also been detected. Hardy *et al.* have cloned a series of cell lines representing each of the various phenotypic forms of NFS-5 (ref. 5). These cloned lines have been designated as follows: 5.3 ($\mu^+ \kappa^-$), 5.4 κ ($\mu^+ \kappa^+$), 5.4 $\kappa\lambda$ ($\mu^+ \kappa^+ \lambda^+$) and 5.4 λ ($\mu^+ \lambda^+$). The two λ -expressing lines, 5.4 $\kappa\lambda$ and 5.4 λ , which represent progeny derived from a single expanded population of 5.4 κ , contain immunoglobulin gene rearrangements not present in either the parental lymphoma or the $\mu^+ \kappa^+$ precursor, 5.4 κ . Thus, although all lines contain shared rearrangements on both κ alleles, lines differ in the organization of both the λ and H-chain genes. Distinct rearrangements at the λ locus are found in 5.4 $\kappa\lambda$ and

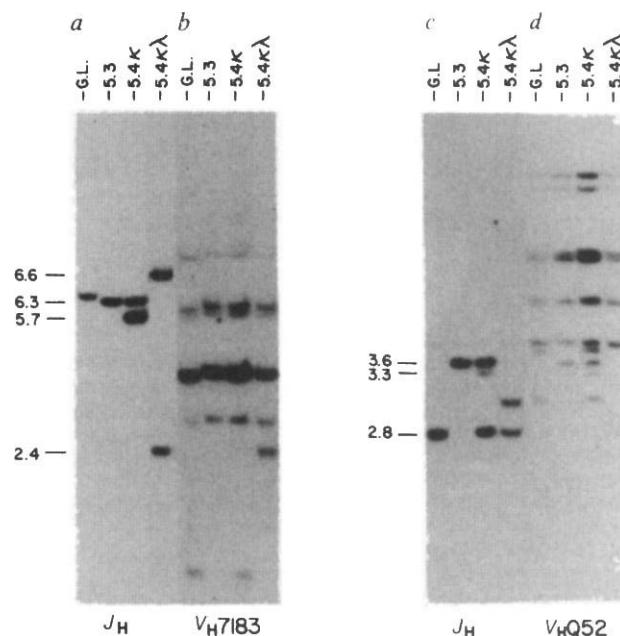


Fig. 1 Context of immunoglobulin variable (V_H) and joining (J_H) gene segments in NFS-5-derived cell lines. Approximately 10 μ g of genomic DNA was digested with either *Eco*RI (a, b) or *Hind*III (c, d), fractionated on 0.8% agarose gels and transferred to nitrocellulose filters²⁷; duplicate blots were assayed for hybridization to the ³²P-labelled J_H probe, pJ11 (a, c)²⁸, and to either a V_H7183 (b) or V_HQ52 (d)²⁹ gene probe. After hybridization, blots were first washed at low stringency ($3 \times$ SSC, 68°C) and then at high stringency ($0.2 \times$ SSC, 68°C). The cell lines assayed include: 5.3 ($\mu^+ \kappa^-$), 5.4 κ ($\mu^+ \kappa^+$) and 5.4 $\kappa\lambda$ ($\mu^+ \kappa^+ \lambda^+$). G.L., germline DNA from a NFS/N $\times$ NZB RI line homozygous for the NFS/N heavy-chain locus. The J_H region probe, pJ11, consists of a 2.0-kb *Bam*HI/*Eco*RI fragment containing the J_H3 and J_H4 gene segments subcloned into pBR322; the V_HQ52 probe (pV_HQ52NHhaI) consists of a 0.3-kb *Hha*I fragment subcloned into the *Sma*I site of pUC12; the V_H7183 probe (pV_HSAPC-15) derives from a 0.9-kb *Eco*RI/*Hae*III restriction fragment subcloned into the *Eco*RI/*Sma*I site of pUC12. The 4.1-kb fragment seen in all lanes of c and d is due to plasmid contamination. Sizes (in kb) of restriction fragments hybridizing to both V_H and J_H probes are given on the left of the autoradiographs. The 3.3-kb fragment in c represents a nonproductive V_HDJ_H2 rearrangement (see text, and data not shown). The light hybridization of this band is thought to be due to a *Hind*III site between the J_H3 and J_H4 gene segments which cuts the pJ11-hybridizable region into two fragments of 2.8 and 3.3 kb; the 3.3-kb fragment contains the variable region and 400 bases of sequence hybridizable to pJ11.

5.4 λ ; whereas 5.4 $\kappa\lambda$ contains a $V_{H1}J_{H1}$ gene, 5.4 λ has a rearranged $V_{H2}J_{H2}$ gene⁵. As these rearrangements are not evident in either the parental lymphoma or 5.4 κ , λ light-chain gene recombination has occurred during growth *in vitro*. In addition, these λ -expressing lines are unusual in that they contain H-chain gene rearrangements not present in the $\mu^+ \kappa^+$ precursor, 5.4 κ . As both H-chain alleles in 5.4 κ consist of complete V_HDJ_H variable regions, these 'secondary' H-chain gene rearrangements have been investigated further.

DNA of high relative molecular mass (M_r) from NFS-5.3, 5.4 κ and 5.4 $\kappa\lambda$ was digested with either *Eco*RI or *Hind*III and analysed by Southern blot hybridization using the J_H -region probe, pJ11 (Fig. 1a, c). Both H-chain alleles of 5.4 κ are rearranged and detected in *Hind*III digests as 3.6-kilobase (kb) and 3.3-kb restriction fragments (Fig. 1c). Restriction fragments of identical size are detected using a V_HQ52 gene probe (Fig. 1d), suggesting that these rearrangements consist of complete V_HDJ_H joins which utilize V_H genes belonging to the V_HQ52 gene family. This conclusion is supported by Northern

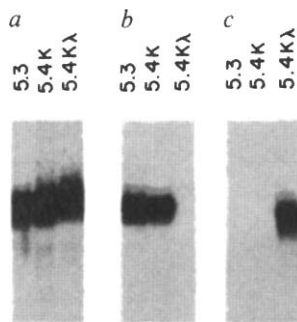


Fig. 2 Northern blot analysis of total RNA from NFS-5-derived cell lines using C_μ (a), V_HQ52 (b) and V_H7183 (c) gene probes. Total cellular RNA was prepared from homogenized cells lysed in 6 M urea/3 M LiCl, extracted with phenol/chloroform and ethanol-precipitated³⁰. Approximately 10 μ g of RNA from each cell line was electrophoresed through a 1.5% agarose/formaldehyde gel³¹, transferred to nitrocellulose and hybridized with nick-translated ³²P-labelled probe in 50% formamide, 5 $\times$ SSC, 1 $\times$ Denhardt's, 20 mM NaHPO₄ pH 6.5, 10% dextran sulphate at 42 $^\circ$ C (ref. 32). Blots were initially washed at low stringency (2 $\times$ SSC, 0.1% SDS, 52 $^\circ$ C), followed by high-stringency washes (0.1 $\times$ SSC, 0.1% SDS, 52 $^\circ$ C). Transcript sizes correspond to those expected for mature μ mRNA.

blot analysis which indicates that 5.4 κ synthesizes a mature transcript that hybridizes to both C_μ (Fig. 2a) and V_HQ52 (Fig. 2b) gene probes. This transcript appears to be encoded by the 3.6-kb *Hind*III restriction fragment, as the $\mu^+ \kappa^-$ cell line, 5.3, which has retained this restriction fragment but deleted the 3.3-kb fragment (Fig. 1c), nevertheless continues to produce both mature μ messenger RNA (Fig. 2a) and protein¹⁸. We conclude that the 3.6-kb *Hind*III fragment contains the productive H-chain allele. This allele is rearranged to J_H4 because the 5.3 cell line, which is haploid at the J_H locus (see Fig. 1a, c), has deleted the 2.8-kb germline-encoded *Hind*III fragment; this deletion is diagnostic of rearrangements to J_H4 (Fig. 1c). The 3.3-kb *Hind*III fragment contains a J_H2 -rearranged H-chain allele, as determined both by hybridization using J_H -specific probes and detailed restriction mapping of the cloned allele (R.K., unpublished data). This allele appears to be nonproductive as J_H2 -containing transcripts cannot be detected in 5.4 κ RNA (data not shown).

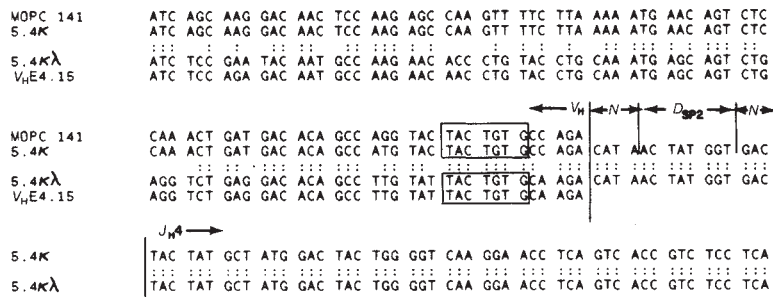
Continued rearrangement of both the productive and nonproductive H-chain alleles is evident in 5.4 $\kappa\lambda$. The relationship between these alleles has been determined by Southern blot analysis using J_H1 - J_H2 and J_H3 - J_H4 -specific probes (data not shown). Rearrangement of the nonproductive allele results in a

change in the *Hind*III fragment size, from 3.3 to 2.8 kb (Fig. 1c). The 2.8-kb J_H -hybridizable fragment is also detected using the V_HQ52 probe (Fig. 1d), indicating that this allele still contains a complete V_HQ52DJ_H rearrangement. Accompanying this rearrangement, germline genes in both the V_HQ52 (Fig. 1d) and V_H7183 (Fig. 1b) gene families have been deleted suggesting an event occurring over many kilobases of DNA extending from the J_H region to the germline V_H locus. Rearrangement of the productive allele in 5.4 $\kappa\lambda$ is accompanied by a change in the *Hind*III fragment from 3.6 to 3.0 kb (Fig. 1c), and a change in the *Eco*RI fragment from 6.3 to 2.4 kb (Fig. 1a). As shown in Fig. 1d, the 3.0-kb fragment does not hybridize to the V_HQ52 gene probe, despite its derivation from the 3.6-kb V_HQ52 -containing variable region (see above). Rather, as shown in the *Eco*RI digests, the productive allele of 5.4 $\kappa\lambda$, contained on a 2.4-kb fragment (Fig. 1a), hybridizes to a V_H7183 gene probe. Thus, secondary rearrangement of the productive allele results in replacement of the initial V_HQ52 gene by a V_H7183 gene. Northern blot analysis of total RNA from 5.4 $\kappa\lambda$ indicates that this replacement is accompanied by a change in V_H RNA expression from V_HQ52 to V_H7183 (Fig. 2c).

To gain insight into the mechanism of this V_H gene replacement, we have sequenced the expressed variable regions of 5.4 κ and 5.4 $\kappa\lambda$ using RNA primer extension techniques¹⁹. Analysis of 5.4 κ yields a single V_HQ52DJ_H4 sequence (Fig. 3), indicating that the remaining V_HQ52 rearrangement (the 5.7-kb *Eco*RI fragment in Fig. 1) in this cell line does not accumulate RNA to any significant extent and is therefore nonproductive at the phenotypic level. Analysis of the H-chain mRNA in 5.4 $\kappa\lambda$ similarly yields a single $V_H7183DJ_H4$ sequence. No V_HQ52 transcript is detected in these cells by Northern analysis (Fig. 2b), indicating that the remaining V_HQ52 rearrangement (the 6.6-kb *Eco*RI fragment) is not expressed. As the productive V_H7183 rearrangement in 5.4 $\kappa\lambda$ derives from a V_HQ52 rearrangement, V_H gene replacement has occurred during propagation *in vitro*. This conclusion is supported by sequence identity of the N , D_H and J_H segments expressed in 5.4 κ and 5.4 $\kappa\lambda$. The N sequences are of particular interest as they are created *de novo* during the V - D recombinational process. Identity of N sequences confirms the common origin of these alleles (see Fig. 3). The variable gene segments of these alleles, however, show marked differences in sequence. Comparison of the 5.4 κ sequence with a representative germline V_HQ52 sequence (M141; ref. 16) confirms that this gene is a member of the V_HQ52 gene family. A similar comparison of the 5.4 $\kappa\lambda$ sequence with a representative germline V_H7183 sequence (V_H E4.30; ref. 15) establishes that the productive allele of 5.4 $\kappa\lambda$ is derived from a member of the V_H7183 gene family. Recombination, therefore, results in replacement of most or all of the coding sequence of

Fig. 3 Nucleotide sequence comparison between the expressed alleles of 5.4 κ and 5.4 $\kappa\lambda$. Sequence homologies are shown and reference V_HQ52 (MOPC 141)¹⁶ and V_H7183 (V_H E4.15)¹⁵ germline sequences are given to indicate V_H gene family relationships (>80% sequence homology)³⁰. Because of identity in the terminal three bases of the V_H segments, the exact site of the recombination cannot be determined precisely; however, four bases 5' of the V_HD_H junction, a single germline-encoded base difference between 5.4 κ (cytosine) and 5.4 $\kappa\lambda$ (adenine) indicates the 5' boundary of the recombination. Base sequences identical to the recombination signals found 5' of many D elements (5'TACTGTG 3')¹⁷ are boxed.

Methods. Sequences were obtained using RNA primer extension techniques according to Shlomchik *et al.*¹⁹. Briefly, total cellular RNA was extracted in guanidinium isothiocyanate³³, and poly(A)⁺ RNA was selected on oligo-(dT)-cellulose columns³². Synthesis of complementary DNA was primed with 50 ng of the 5'-end-labelled oligonucleotide (5'GCAGGAGAGGAGGGGA 3') homologous to μ constant-region exon 1. Labelled primer was annealed to 80–120 μ g of poly(A)⁺ RNA, and reverse transcription carried out in 100 mM Tris(8.3), 140 mM KCl, 10 mM MgCl₂, 0.5 mM dNTPs, 10 mM dithiothreitol, 120 U RNasin, 40 U reverse transcriptase for 2 h at 42 $^\circ$ C. The cDNA was obtained by fractionation on a 5% polyacrylamide/7M urea gel for 2–3 h at 30 V cm⁻¹. Gel slices were chopped and eluted overnight at 37 $^\circ$ C in 0.5 M NH₄Ac, 1 mM EDTA. Full-length cDNA was ethanol-precipitated and washed extensively before sequencing by modified chemical degradations, as described by Rubin and Schmid³⁴ and Bencini *et al.*³⁵.



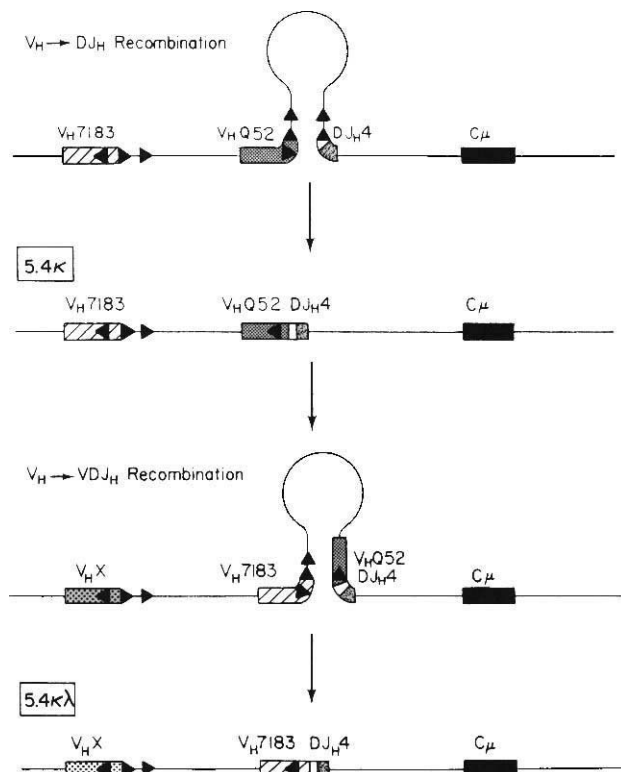


Fig. 4 Model of the V_H gene recombinations leading to the formation of the expressed H-chain allele of 5.4 $\kappa\lambda$. Heptameric and nanomeric signal sequences are indicated by black arrows, and component gene segments are indicated by patterned overlays. Both N and D_{SP2} sequences are depicted together in the figure as an open box. Top panel shows the initial V_HQ52 to DJ_H4 recombination, involving interaction of signal sequences found 3' of the V_HQ52 gene with those 5' of the DJ_H4 complex to form a hypothetical stem-loop structure. Excision at the base of the stem and fusion of V_HQ52 gene to DJ_H4 completes the formation of the intact variable region of 5.4 κ . Subsequent recombination between an upstream germline V_H7183 gene segment and the productive V_HQ52 rearrangement is postulated to occur via interaction of the heptameric and nonameric signal sequences 3' of the V_H7183 gene with the embedded heptamer of the rearranged V_HQ52 gene, resulting in the replacement of the V_HQ52 coding region by a V_H7183 gene segment. This recombination is precise, and does not result in the addition or deletion of bases: the 5' portion of the newly formed variable region is contributed by the incoming V_H7183 gene, whereas the 3' portion, consisting of N , D_{SP2} and J_H4 sequences, has been retained from the recombinations which initially formed the allele.

the resident V_HQ52 gene with that of the incoming V_H7183 gene. This recombination has occurred so as to maintain a continuous open reading frame. Because of the identity of the 3'-terminal three bases of the germline V_H gene segments (both gene families use an AGA for this codon), the precise 3' point of the recombination cannot be identified. However, a single germline-encoded base difference between 5.4 κ and 5.4 $\kappa\lambda$ delineates the 5' boundary of the recombination four bases 5' to the V - D junction.

A number of mechanisms such as gene conversion⁴ and homologous recombination²⁰ may be invoked to explain this V_H gene replacement. We consider yet a third model of V_H - V_H recombination by a mechanism analogous to V_H - D_H recombination. Embedded in the V_H coding region, close to the site of the recombination, is the sequence 5' TACTGTG 3'²¹, which is identical to the signal heptamer found 5' of many D_H gene segments¹⁷. We suggest that this heptamer functions as a pairing element in mediating V_H - V_H recombination (see Fig. 4). Here

a V_HQ52 gene is paired by its 3' signal sequences to the signal sequences found 5' of DJ_H4 ; shown as a stem-loop structure. Endonucleolytic cleavage at the base of the stem, and fusion of the V_HQ52 gene segment to this DJ_H4 intermediate results in the formation of the productive allele found in 5.4 κ . Further rearrangement of this allele is postulated to occur by an interaction between the embedded heptamer of the V_HQ52 gene and the signal sequences 3' of one of the germline V_H7183 genes still present on the chromosome. This results in the replacement of the V_HQ52 gene by a V_H7183 gene to give the productive allele found in 5.4 $\kappa\lambda$.

Although conventional immunoglobulin gene recombination is correlated with the presence of both heptameric and nonameric signal sequences, recombination at isolated heptamers has also been reported. A precedent for heptamer-mediated recombination has been obtained at the κ locus, where rearrangements of both V_κ and J_κ gene segments to an isolated heptamer in the J_κ - C_κ intron have been reported^{20,22}. Moreover, in many λ -producing plasmacytomas and hybridomas, deletion of the C_κ gene segment is mediated by a recombinational event between this intron heptamer and heptamer-nonamer sequences located 3' of the C_κ locus (such C_κ deletion has not, however, been observed in 5.4 $\kappa\lambda$)^{23,24}. It has been suggested that such a deletion may be an important preamble to λ gene recombination²³, implying that these heptamer-mediated recombinations may have biologically important consequences.

Yancopoulos *et al.*¹⁵ and Perlmutter *et al.*²⁵ have recently described the preferential usage of V_H genes from the J -most proximal V_H gene family early in fetal development. This preference is not evident in the adult repertoire where the V_H gene families appear to be used at frequencies reflecting the relative complexity of each gene family²⁶. We suggest that V_H gene replacement mediated by the embedded heptamer may represent one mechanism leading to the random V_H usage observed in the adult repertoire. Indeed, the strong conservation of this heptameric sequence within the FR3 region where it encodes Tyr and Cys residues at positions 91 and 92 supports this hypothesis: over 70% of V_H genes which have been examined, including all functional members of the J -proximal V_H families (for example, V_H7183 , V_HQ52 and V_HS107)⁷⁻¹⁷, have been found to contain this heptameric sequence. Thus, conservation is evident not only at the amino-acid sequence level, but also at the nucleotide level where little, if any, third base degeneracy is observed at both the Tyr and Cys codons. Such a high degree of conservation, even among highly diverged V_H gene families, suggests that this heptameric sequence has an important role in the somatic diversification of the antibody repertoire.

We thank G. Lennon and M. Shapiro for useful discussions, J. Brill Dashoff for expert technical assistance and M. Shlomchik for sequencing reagents. This work was supported in part by NIH grants GM-20964-13, CA-04681 and HD01287. R.R.H. was supported by the American Cancer Society, California Division (Fellowship S-12-84).

Received 6 May; accepted 28 May 1986.

1. Yancopoulos, G. & Alt, F. W. *A. Rev. Immun.* **4**, 339-368 (1986).
2. Weigert, M., Cesari, I. M., Yonkovich, S. J. & Cohn, M. *Nature* **228**, 1045-1047 (1970).
3. McKean, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3180-3184 (1984).
4. Dildrop, R., Brüggemann, M., Radbruch, A., Rajewski, R. & Beyreuther, K. *EMBO J.* **5**, 635-640 (1982).
5. Hardy, R. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 1438-1442 (1986).
6. Kurosawa, Y. & Tonegawa, S. *J. exp. Med.* **155**, 201-218 (1982).
7. Bothwell, A. L. *et al. Cell* **24**, 625-637 (1981).
8. Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* **25**, 59-66 (1981).
9. Givol, D. *et al. Nature* **292**, 426-430 (1981).
10. Kaartinen, M., Griffiths, G. M., Martham, A. F. & Milstein, C. *Nature* **304**, 320-324 (1981).
11. Loh, D. Y., Bothwell, A. L. M., White-Schaff, M. E., Imanishi-Kari, T. & Baltimore, D. *Cell* **33**, 85-93 (1983).
12. Near, R. I. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2167-2171 (1984).
13. Sims, J. *et al. Science* **216**, 309-311 (1982).
14. Kaartinen, M. *et al. J. Immun.* **130**, 937-945 (1983).
15. Yancopoulos, G. D. *et al. Nature* **311**, 727-733 (1984).
16. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
17. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981-992 (1980).
18. Davidson, W. F. *et al. J. Immun.* **133**, 744-753 (1984).

19. Shlomchik, M. J. *et al. J. exp. Med.* (in the press).
20. Seidman, J. G. & Leder, P. *Nature* **286**, 779-783 (1980).
21. Early, P., Nottenburg, C., Weissman, I. & Hood, L. *Molec. cell. Biol.* **2**, 829-836 (1982).
22. Kelley, D. E. *et al. Molec. cell Biol.* **5**, 1660-1675 (1985).
23. Durdik, J., Moore, M. W. & Selsing, E. *Nature* **307**, 749-752 (1984).
24. Siminovitch, K. A., Bakshi, A., Goldman, P. & Korsmeyer, S. J. *Nature* **316**, 260-262 (1985).
25. Perlmutter, R. M., Kearney, J. F., Chang, A. P. & Hood, L. E. *Science* **227**, 1597-1600 (1985).
26. Dildrop, R., Krawinkel, V., Winter, E. & Rajewsky, K. *Eur. J. Immun.* **15**, 1154-1156 (1985).
27. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
28. Marcu, K. B., Banerji, J., Penncavage, N. A., Long, R. & Arnheim, N. *Cell* **22**, 187-196 (1980).
29. Brodeur, P., Riblet, R. *Eur. J. Immun.* **14**, 922-930 (1984).
30. Aufray, C. & Rougeau, F. *Eur. J. Biochem.* **107**, 303-314 (1980).
31. Thiele, C. J., Reynolds, C. P. & Israel, M. A. *Nature* **313**, 404-406 (1985).
32. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
33. Chirgwin, J., Przybyla, A., MacDonald, R. & Rutter, W. *Biochemistry* **18**, 5294-5299 (1979).
34. Rubin, C. & Schmid, C. *Nucleic Acids Res.* **8**, 4613-4619 (1980).
35. Bencini, D., O'Donovan, G. & Wild, J. *Biotechniques* **4** (January/February 1984).

Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer

Judah Weinberger*†‡, David Baltimore†‡
& Phillip A. Sharp*†

* Center for Cancer Research and † Department of Biology,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

‡ Whitehead Institute for Biomedical Research, Cambridge,
Massachusetts 02142, USA

The intron separating the variable- and constant-region exons of the rearranged immunoglobulin heavy-chain locus contains a lymphocyte-specific transcriptional enhancer^{1,2}. The enhancer is a member of a class of *cis*-acting, tissue-specific, transcriptional control elements which are characterized by orientation-independent and relatively position-independent function^{1,2}. *In vivo* analysis of the position of DNA-binding factors, by assessing the availability of specific bases to chemical modification, has identified four sequence clusters within the heavy-chain enhancer, denoted E1 to E4 (refs 3, 4). These sites are protected (that is, occupied) only in B lymphocytes. A consensus sequence relationship (consensus CAGGTGGC) between these four sites was suggested where three of the sites conformed to the consensus in seven of eight positions while the other was homologous in six of eight positions. We proposed that a single *trans*-acting factor might recognize all four sites^{3,4}. Using an assay involving gel electrophoresis of DNA-protein complexes⁵⁻⁸ to detect sequence-specific DNA binding factors that recognize these related motifs, we have now identified a mouse B-cell nuclear factor (NF- μ E1) which binds specifically to one such motif within the mouse heavy-chain gene enhancer. This factor binds poorly, if at all, to the other related motifs, and other factors have been identified which interact preferentially with some of these latter motifs. Dimethyl sulphate interference experiments suggest that the NF- μ E1 factor is in contact with at least the guanine residues in the sequence GATGGCCGATC. This factor seems to be present in both lymphoid and non-lymphoid cell lines.

First, we searched for enhancer-binding factors for these sequence clusters in nuclear extracts from B-cell lines. For this we used a gel electrophoresis DNA-binding assay⁵⁻⁸ in which DNA-protein complexes are resolved by their slower mobility in a low-ionic-strength polyacrylamide gel. A 221-base pair (bp) *Hinf*I fragment containing all *in vivo* identified binding sites was subcloned into the *Sma*I site of the pUC-13 vector (Fig. 1). The *Hind*III-*Pvu*II fragment containing the E1 site and a segment of the polylinker (henceforth denoted as the E1 probe)

was 3'-end-labelled and then incubated with increasing amounts of nuclear extract derived from a mouse B-cell lymphoma line, WEHI-231. The duplex alternating copolymer poly(dI-dC)·poly(dI-dC) was included in the binding reaction to suppress the effect of sequence-nonspecific binding of proteins. A distinct band, a putative protein-DNA complex, with a slower electrophoretic mobility than the free fragment, was observed (Fig. 2a). At higher concentrations of input nuclear extract, additional distinct species with even further retarded mobilities were noted. Competition experiments showed that these species resulted from non-sequence-specific complex formation (data not shown). To test the sequence-specific nature of the major complex, we examined the ability of DNA fragments containing the E1 sequence or other sequences to compete for factor binding to the E1 probe. Two restriction fragments were derived from the mouse immunoglobulin heavy-chain enhancer region. The 400-bp *Xba*I-*Pst*I fragment (μ 400), which contains E1 sequences (Fig. 1), efficiently competed with radiolabelled E1 probe for binding (Fig. 2b). A 300-bp *Pst*I-*Eco*RI fragment (μ 300) containing the balance of the enhancer (Fig. 1), including a partial E2 site and complete E3 and E4 sequences, was, however, unable to compete for NF- μ E1 binding of the test fragment. Molar ratios of competitor DNA to probe DNA of 20-50 did not reduce the level of specific binding. Thus, there is no high-affinity NF- μ E1-binding site in the μ 300 fragment.

The μ 300 fragment is recognized by two distinct factors which generate distinct mobility complexes in the gel shift assay^{8,9}. One of these, IgNF-A, recognizes an octamer motif, ATTTGCAT, conserved between the mouse immunoglobulin heavy- and light-chain promoters, as well as the heavy-chain enhancer⁸. Another sequence-specific factor binds to a 70-bp *Alu*I fragment containing E3⁹. A possible way in which the four binding sites occupied *in vivo* might result from binding of a single factor, and yet without DNA containing the E3 site competing for binding to the E1 site, would be if binding to the E3 site was of lower affinity. To assess this possibility, a 5'-end-labelled E3 probe, the 70-bp *Alu*I fragment, was incubated in WEHI-231 nuclear extract together with either μ 400, containing the E1 site, or μ 300, containing the balance of the known *in vivo* binding sites. Figure 2b shows that binding to the E3 probe was not efficiently competed for by excess E1 binding sites while the same binding was efficiently competed for by homologous sequences in the μ 300 fragment. Titrations of low molar ratios of μ 400 and μ 300 fragments compete for E1 and E3 probe

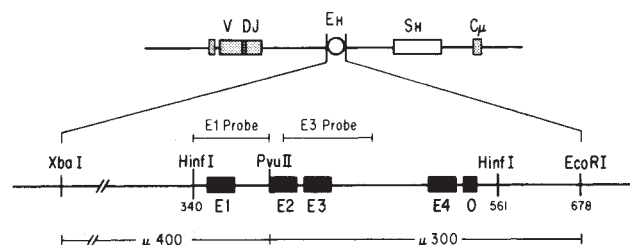
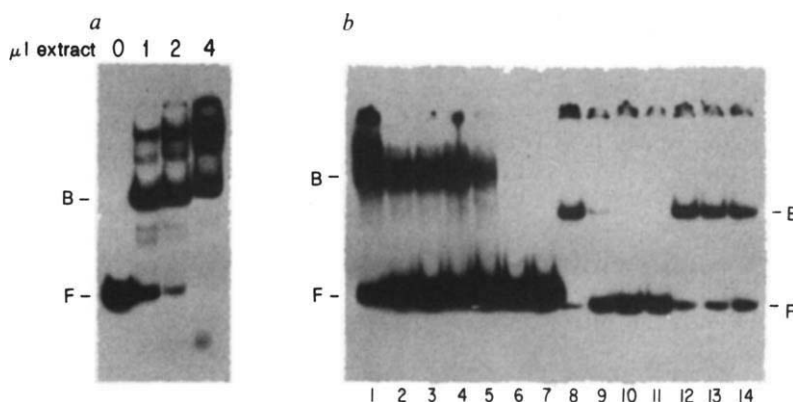


Fig. 1 Scheme of probes and competitor fragments used. The immunoglobulin heavy-chain locus is shown with the intron containing the enhancer (open circle) and the μ switch region (S_{μ} , open bar) separating the rearranged V-D-J (variable-diversity-joining) exons and C_{μ} , the constant-region μ -chain gene. The enhancer is located in a 678-bp *Xba*I-*Eco*RI fragment with the indicated restriction sites. E1-E4 and O refer to sequences that are protected from chemical modification in intact B cells³ or in their nuclei⁴. The *Hinf*I fragment, containing all the indicated sites, was blunted and subcloned into the *Sma*I site of pUC-13. The E1 probe was derived by cutting the resulting plasmid with *Hind*III, in the polylinker, and *Pvu*II. The E3 probe is the 70-bp *Alu*I fragment spanning the indicated region. A *Pst*I site used for deriving two fragments indicated on the lowest line, μ 400 and μ 300, is located 6 bp downstream of the *Pvu*II recognition site.

Fig. 2 Binding of a nuclear factor to E1 and competition analysis. *a*, Binding in WEHI-231 nuclear extract. The E1 probe was incubated with nuclear extract as described below. The bound protein-DNA complex and free DNA fragment were resolved on a polyacrylamide gel. The free fragment is labelled F, and a bound complex, B, *b*, Competition analysis of binding to E1 and E3. Lanes 1-7, binding reactions using 1 ng of a [5'- α -³²P]-end-labelled 70-bp E3 probe. Lane 1, a control reaction in the absence of additional competitor DNA; Lanes 2-4, binding reactions carried out in the presence of 25, 75 or 150 ng, respectively, of the μ 400 fragment (Fig. 1). Lanes 5-7, binding reactions carried out in the presence of 25, 75 and 150 ng of the μ 300 fragment. Lanes 8-14, binding reactions using 1 ng of a radiolabelled 65-bp E1 probe. Lane 8, a control reaction carried out in the absence of additional competitor DNA. Lanes 9-11, binding reactions carried out in the presence of 25, 75 and 150 ng, respectively, of the μ 400 fragment. Lanes 12-14, binding reactions carried out in the presence of 25, 75 and 150 ng, respectively, of the μ 300 fragment.

Methods. The E1 probe was excised as described for Fig. 1, end labelled with [α -³²P]dATP and the large fragment of *Escherichia coli* DNA polymerase I, and isolated by polyacrylamide gel electrophoresis. The ³²P-labelled fragment (~0.5 ng, 10,000 c.p.m.) was incubated with the nuclear extract¹⁴ of a murine B-cell lymphoma, WEHI-231. Binding reactions (10 μ l) contained 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol, 1,600 ng poly(dI-dC)·poly(dI-dC), and the indicated amounts of nuclear extract. After incubation at room temperature for 10 min, the complexes were resolved on a 4% polyacrylamide gel (acrylamide:bisacrylamide weight ratio of 29:1) in 45 mM Tris-borate, 1 mM EDTA pH 8.0 buffer. The gel was dried and autoradiographed.



binding with the expected quantitation (data not shown). Thus, NF- μ E1 and NF- μ E3⁹ represent distinct factors with high affinity for distinct cognate sequences. Similar competition experiments using the human immunoglobulin heavy-chain enhancer¹⁰ demonstrate the existence of NF- μ E1- and NF- μ E3⁹-binding sites in this fragment. No binding activity specific for E4 or E2 sequences has been detected; however, all the probes

used to date have been derived from fragments that were cut at the *Pst*I site, which may interfere with binding.

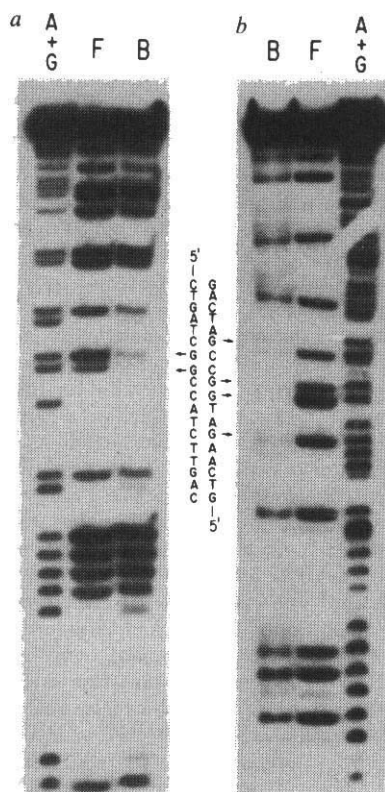
To elucidate the critical contact points for the binding of NF- μ E1 protein, DNA methylation interference analysis was performed¹¹. Dimethylsulphate (DMS) methylates the N-7 position of guanine which extends into the major groove of the B-helix¹². The E1 probe was partially methylated with DMS, a binding reaction was performed with the WEHI-231 nuclear extract, and bands of the free fragment and the protein-DNA complex were resolved as described in Fig. 2. Fragments recovered from these bands were cleaved at methylated guanine residues with piperidine and the cleavage products were displayed on denaturing sequencing gels. Positions at which methylation of guanine residues interferes with factor binding should not be cleaved in fragments recovered from the DNA-protein complex band. Comparison of the cleavage pattern of radioactive fragments from the two bands revealed several discrete sites on both strands of the E1 probe where methylation interfered with binding (Fig. 3). These data suggest that the binding site of the factor must include contacts with the guanine residues in the sequence $\begin{smallmatrix} \text{GATGGCCG} \\ \text{CTACCGGC} \end{smallmatrix}$. The identified interfering guanine residues are precisely the residues protected against DMS methylation *in vivo*³. Delineation of contact residues for NF- μ E1 binding underscores a lack of significant homology with other putative enhancer-binding sites. In particular, the sequences identified for E3 binding^{3,4}, TGTGGCAAG, differ from the E1-binding site by four of eight bases. Furthermore, the above methylation pattern of the E1 site suggests a recognition sequence that only partially overlaps the consensus homology previously deduced by comparison of the four *in vivo* binding sites.

When nuclear extracts from a panel of pre-B cells, B cells, plasma cells, T cells and various non-lymphoid cells were assessed for the level of NF- μ E1 activity (Fig. 4), all were found to contain an activity that generated a band with the E1 probe which co-migrated with the NF- μ E1 band from WEHI-231 cells. Some cell lines showed evidence that minor species migrated more rapidly than the NF- μ E1 band. These complexes may represent degradation products of NF- μ E1 because repeated freezing and thawing of some nuclear extracts increased the levels. Competition experiments and methylation interference analysis performed in various non-lymphoid nuclear extracts revealed sequence specificity identical to that described above for B-cell nuclear extracts (data not shown).

The biological function of NF- μ E1 and its biochemical mode

Fig. 3 Methylation interference analysis of binding to E1. The E1 probe was uniquely radiolabelled at the *Hind*III site in the polylinker (5' to *Hin*II site of the enhancer, see Fig. 1) by the large fragment of *E. coli* DNA polymerase I (*a*) or by polynucleotide kinase (*b*). F and B refer, respectively, to DNA recovered from the free and bound species, the products of the binding reaction. A+G chemical cleavage ladders¹² of the E1 probe were co-electrophoresed to map the binding domain.

Methods. The probe (5 ng, 10⁵ c.p.m.) was methylated with DMS for 8 min at 20 °C¹². After ethanol precipitation, a 30- μ l binding reaction was carried out using the methylated probe, 6 μ l of WEHI-231 nuclear extract and 3,600 μ g of poly(dI-dC)·poly(dI-dC). The remainder of the reaction conditions were as described in the Fig. 2a legend. The products of the reaction were resolved as in the Fig. 2a legend. After autoradiography to visualize the various species, DNA was eluted from the free and bound fragments, and cleaved with 1 M piperidine in 100 μ l at 90 °C for 30 min¹². After overnight lyophilization, the products were analysed by separation in an 8% polyacrylamide gel (19:1) in the presence of 8 M urea followed by autoradiography at -70 °C with an intensifying screen. Essentially the entire probe is displayed in this figure and the only interference sites detected are those indicated.



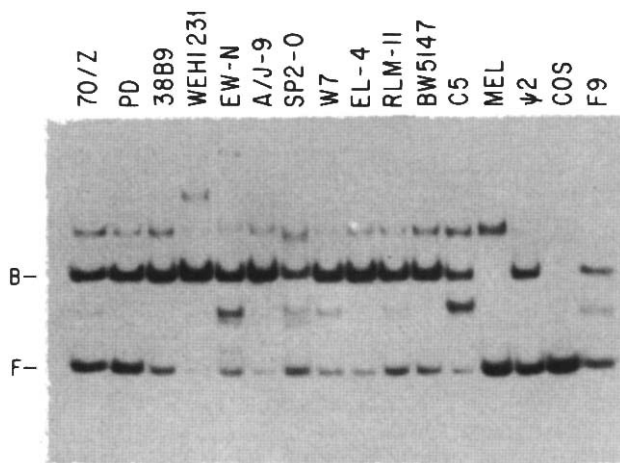


Fig. 4 Analysis of distribution of E1-binding activity in various cell types. Nuclear extracts were prepared¹⁴ from murine or human pre-B-cell lines (70/Z, PD, 38B9), B-cell lines (WEHI-231, EW-N), a myeloma (W7), T-cell lines (EL-4, RLM-11, BW-5147), a primitive B-cell progenitor cell line¹⁵ (C5, a derivative of Ba/F3, a line like those described in ref. 15), an erythroid line (MEL), a fibroblast line (Ψ 2), simian virus 40-transformed African green monkey kidney cell line (COS), and an embryonal cell line (F9). Binding reactions were carried out with the E1 probe as described in Fig. 2 legend. The results of the reaction were analysed by electrophoresis on native polyacrylamide gels. Autoradiography of the dried gels was performed at -70°C with intensifying screens. Longer exposures of the autoradiogram reveal detectable levels of a species co-migrating with B, the bound complex, in MEL and COS cells.

of action are unknown. The evolutionary conservation of an E1-binding site between mouse and human immunoglobulin genes suggests an important biological function. Deletions of sequences containing the E1 site have been shown to produce severalfold reduction in the transcriptional stimulation activity of the immunoglobulin enhancer element in B cells^{2,13}. The E1 element may be a transcriptional control element, with either constitutive or lymphocyte-restricted transcriptional activation. Perhaps the most direct evidence for the involvement of binding to the E1 site is the previous *in vivo* methylation protection analysis. A factor is bound to the E1 site *in situ* which forms contacts with identical guanine residues to those found with the solubilized NF- μ E1 factor³. This suggests that NF- μ E1 is bound to the E1 site in B lymphocytes.

In the above studies, NF- μ E1 binding *in vivo* was found only in cells of the B lineage. *In vitro*, however, factors with identical binding specificity are present in extracts from cells of non-lymphocytic lineage. Although similar binding specificities are defined, the proteins found in the various cell types may differ in important aspects that would not be detected by this assay. An obvious explanation for these apparently inconsistent observations is that the immunoglobulin enhancer is in an inactive conformation in non-B cells and thus inaccessible for binding by NF- μ E1. The ubiquitous expression of an NF- μ E1 activity suggests that this factor is probably important for expression of genes in non-lymphocytic cells. It is possible, for example, that NF- μ E1 binding has completely different consequences in cells of different developmental lineages or stages or in different genomic contexts, including *cis*-dominant transcriptional suppression. These possibilities are being investigated.

Several sites (μ E2, μ E3 and μ E4) with limited sequence homology to the E1-binding site have been identified within the immunoglobulin heavy-chain enhancer. It has been suggested that these four sites bind a common factor. Evidence presented here, in fact, suggests that this is not the case. NF- μ E1 clearly does not bind to the μ E3 or μ E4 sites; conversely, a different factor, NF- μ E3, binds with high affinity to E3 but does not bind E1 to a significant degree⁹. Thus, the immunoglobulin heavy-chain enhancer element must be recognized by at least three

and probably more different factors. How the combination of these factors sum to a cell-type-specific enhancer is a fascinating problem.

We thank Dr L. Staudt for nuclear extracts from various cell lines, Dr A. Hayday for a plasmid containing the human heavy-chain enhancer, Drs J. LeBowitz and H. Singh for helpful discussions and a critical reading of the manuscript and M. Siafaca for preparation of the manuscript. J.W. acknowledges support from NIH Clinical Investigator Award HL01633. This work was supported by grants from NIH (CA38660 and GM32467), NSF (DCB8502718) and partially from NCI Cancer Center Support (Core) (CA14051) to P.A.S. and by a grant from the American Cancer Society (IM-355Q) to D.B.

Received 31 March; accepted 6 June 1986.

1. Banerji, T., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
2. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-728 (1983).
3. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. *Science* **227**, 134-140 (1985).
4. Church, G. M., Ephrussi, A., Gilbert, W. & Tonegawa, S. *Nature* **313**, 798-801 (1985).
5. Fried, M. & Crothers, D. M. *Nucleic Acids Res.* **9**, 6505-6525 (1981).
6. Garner, M. M. & Revzin, A. *Nucleic Acids Res.* **9**, 3047-3060 (1981).
7. Carthew, R. W., Chodosh, L. A. & Sharp, P. A. *Cell* **43**, 439-448 (1985).
8. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. *Nature* **319**, 154-158 (1986).
9. Sen, R. & Baltimore, D. *Cell* (in the press).
10. Hayday, A. C. *et al.* *Nature* **307**, 334-340 (1984).
11. Sienbenlist, U. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 122-126 (1980).
12. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-525 (1980).
13. Grosschedl, R., Constantini, F. & Baltimore, D. *Banbury Rep.* **20**, 187-196 (1985).
14. Dignam, J. D., Lebovitz, R. M. & Roder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
15. Palacios, R. & Steinmetz, M. *Cell* **41**, 727-734 (1985).

Deregulated expression of *c-myc* by murine erythroleukaemia cells prevents differentiation

E. V. Prochownik & J. Kukowska

Department of Pediatrics and the Committee on Cellular and Molecular Biology, C. S. Mott Children's Hospital, University of Michigan School of Medicine, Ann Arbor, Michigan 48109, USA

Friend murine erythroleukaemia (F-MEL) cells are a permanent line of primitive erythroid precursors originally derived from the spleens of mice infected with the Friend strain of murine leukaemia virus¹. F-MEL cells differentiate *in vitro* in response to various chemical inducers^{1,2}. Concomitantly with induction, a biphasic regulation of *c-myc* oncogene transcripts is observed³. Within one hour of the addition of dimethyl sulphoxide (DMSO) or hypoxanthine (Hyp), the levels of *c-myc* transcripts fall dramatically and remain virtually undetectable for the next few hours. Between 8 and 24 hours after induction, *c-myc* transcripts return to pre-induction levels and then decline again between 3 and 5 days as most of the cells undergo terminal differentiation. To explore the potential relationship between *c-myc* expression and F-MEL terminal differentiation, we have investigated here whether reversing the early fluctuations in *c-myc* transcript levels affects the ability of F-MEL cells to differentiate. We therefore constructed an amplifiable plasmid vector containing a full-length mouse *c-myc* complementary DNA and introduced it stably into recipient F-MEL cells. The exogenous *c-myc* sequences are transcribed in F-MEL cells and the transcript levels do not change significantly in response to inducing agents. The net result is continued *c-myc* expression following DMSO or Hyp induction and a complete or partial inhibition of F-MEL differentiation.

Figure 1 shows the plasmid used for transfections. The backbone of this vector has been derived from pSV2neo, in which bacterial neomycin resistance (*neo*) sequences have been replaced with a full-length murine *c-myc* cDNA clone. The simian virus 40 (SV40) early splice and poly(A) sites of pSV2neo have been retained. Downstream of *myc* and in the opposite transcriptional orientation, we have introduced an expressible

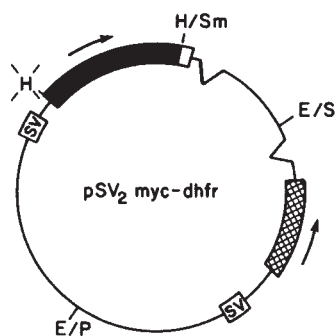


Fig. 1 Structure of an amplifiable mouse *c-myc* expression vector. SV, SV40 promoters; dark box, mouse *myc* cDNA sequences; light box, *neo* 3'-untranslated region sequence; cross-hatched boxes, *dhfr* sequences; V, SV40 splice junctions; H, *Hind*III; Sm, *Sma*I; E, *Eco*RI; S, *Sal*I; P, *Pst*I. **Methods.** All plasmids were purified by the alkaline lysis method²². The plasmid pSV₂neo (ref. 28) was digested with *Hind*III and *Sma*I to remove *neo* sequences and then treated with the large fragment of DNA polymerase I. A 1.5-kilobase (kb) *Hind*III fragment containing the entire coding region of a mouse *c-myc* cDNA was excised from the plasmid pMc-myc 54 (ref. 29), treated with the large fragment of DNA polymerase and ligated into the pSV₂neo-derived vector. The resultant plasmid, with *c-myc* sequences under the control of the SV40 early promoter, was treated with *Eco*RI and the large fragment of DNA polymerase. Into this site we cloned a 2.8-kb *Pst*I-*Sal*I fragment from the plasmid pFR400 (ref. 4) after treating the fragment with T₄ DNA polymerase and the large fragment of DNA polymerase to render it blunt-ended. This fragment contains the SV40 early promoter, a mutant *dhfr* cDNA and splice-polyadenylation sites derived from the gene encoding hepatitis-B virus surface antigen.

full-length cDNA encoding a mutant form of dihydrofolate reductase (*dhfr*) which facilitates a selective amplification of plasmid sequences in the presence of high concentrations of the antimetabolite methotrexate (MTX)⁴. We linearized pSV₂myc-dhfr along with a selectable plasmid marker encoding neomycin resistance. The two plasmids were introduced by electroporation at a 10:1 molar ratio into recipient F-MEL-745 cells. We selected six clones which demonstrated resistance to both neomycin and MTX. Of these, three expressed the exogenously introduced *c-myc* sequences (not shown).

We first examined the levels of endogenous and exogenous *c-myc* transcripts in response to DMSO using an S₁ nuclease protection assay (Fig. 2). Transfected F-MEL cells were first removed from MTX-containing medium for 3–5 days, then washed and cultured in Dulbecco's minimal essential medium (DMEM) containing 1.5% DMSO for varying periods of time. Total RNAs were extracted and used in the S₁ protection assay. The results confirmed the previously reported changes in endogenous *c-myc* transcripts which occur after DMSO addition³. The behaviour of exogenous, pSV₂myc-dhfr-derived *c-myc* transcripts was quite different. Following the addition of DMSO, *c-myc* transcript levels declined by no more than 50%, remaining easily detectable even when endogenous transcripts were not. The net effect was therefore a partial reversal of the early DMSO-induced decline in *c-myc* transcript levels. Of particular interest was the failure of endogenous transcripts in transfected cells to decline significantly by day 5, suggesting that the cells still retained proliferative potential.

The experiments described above indicated that it was possible to reverse at least partially the inhibitory effects of DMSO on *c-myc* transcript levels. We therefore next investigated whether this reversal had any inhibitory effect on the differentiation capacity of F-MEL cells in response to two structurally unrelated inducers of differentiation. F-MEL cells were cultured for varying times in 1.5% DMSO or 5 mM Hyp. The extent of differentiation was monitored daily by determining the fraction of cells that stained with benzidine⁵. The results (Fig. 3) showed that in each of the three clones tested, the extent of benzidine positivity was significantly reduced in relation to non-transfected F-MEL-745 cells or to clones of F-MEL cells which had been

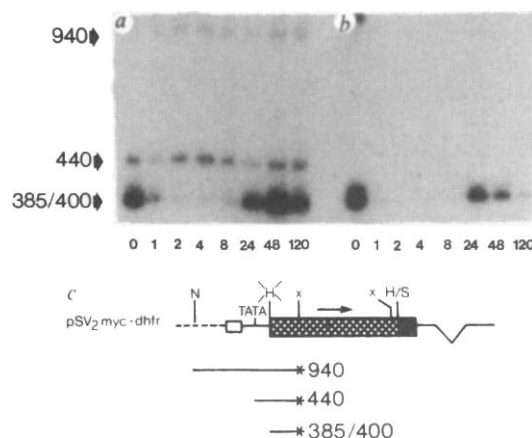


Fig. 2 Levels of mouse *c-myc* transcripts in F-MEL clone 15 (a) or 745 (b) following the addition of DMSO. The dark arrows indicate the 940-nucleotide reannealed input probe, the 440-nucleotide S₁-protected fragment representing pSV₂myc-dhfr-derived exogenous transcripts and the two, unresolved 400- and 385-nucleotide fragments representing endogenous *myc* transcripts³⁰. Numbers below each lane indicate the length of time (in hours) that cells were incubated in the presence of DMSO. c, The S₁ probe was a 940-nucleotide long *Nde*I-*Xho*I fragment, derived from pSV₂ myc-dhfr and end-labelled at the *Xho*I site.

Methods. F-MEL cells (clone 745) were grown in DMEM supplemented with 10% fetal calf serum (FCS), 100 U ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin and 2 mM glutamine (Gibco). Transfections were performed by electroporation as described previously^{31,32} in 1 ml of phosphate-buffered saline containing 1–2 × 10⁷ cells ml⁻¹, 20 µg of linearized SV₂myc-dhfr and 2 µg of linearized pSV₂neo DNA. Immediately after electroporation, cells were kept on ice for 10 min and then replated in a series of 6-well dishes in fresh DMEM containing dialysed FCS. Two days later, the antibiotic G418 (Gibco) was added to a final concentration of 1 mg ml⁻¹. After ~2 weeks, G418-resistant clones were replated in fresh medium containing 0.25 µM MTX (Lederle) plus dialysed FCS. Surviving cells from each well were cloned by limiting dilution in 96-well microtitre dishes. In this way, we insured that each clone originated from an independent transfection event. MTX-resistant clones were expanded with continuous growth in 0.25 µM MTX. Clone 8, 15 and 21 cells were each derived from single-cell isolates of 745 cells stably transfected with pSV₂myc-dhfr and pSV₂neo. For S₁ nuclease studies, clones were grown for 3–5 days in the absence of MTX. DMSO was added for the times indicated and total cellular RNAs were extracted by the guanidine hydrochloride method³³. Aliquots (10 µg) of each RNA were hybridized with an end-labelled 940-nucleotide *Nde*I-*Xho*I fragment derived from pSV₂myc-dhfr (c)³⁴. This probe was labelled with polynucleotide kinase and [γ -³²P]ATP (Amersham) to a specific activity of ~10⁷ d.p.m. µg⁻¹. Hybridization reactions contained 10⁷ d.p.m. of probe and 10 µg of RNA in a total volume of 10 µl of hybridization solution (80% formamide, 300 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA). Reactions were placed at 85 °C for 5 min followed by a 16-h incubation at 51 °C, then treated with 300 U of S₁ nuclease (Sigma) at 37 °C for 1 h, precipitated with isopropanol and subjected to analysis on a 1.8% agarose gel.

transfected with a derivative of the plasmid shown in Fig. 1 in which *myc* sequences were not included or were reversed. Control F-MEL cells demonstrated 85–95% benzidine positivity after 5 days of induction whereas pSV₂myc-dhfr clones showed 10–50% staining.

Transfected clones also failed to accumulate β ^{major}-globin messenger RNA as determined by Northern blot analysis of cellular RNAs (not shown). Furthermore, unlike control F-MEL cells, all three transfected clones continued to proliferate after they had been incubated for 7 days in medium containing 1.5% DMSO (not shown). We conclude from these studies that the perturbations of *myc* expression seen in these clones are associated with the preservation of features normally associated with the uncommitted, non-terminally differentiated state.

The inability of clone 21 cells to differentiate as well as clones 8 and 15 might have been related to levels of exogenous *c-myc* transcripts. Indeed, we found that these levels are two to threefold lower in clone 21 cells than in the other clones (not shown). To address this aspect more directly, we grew cells from clone 21 in increasing concentrations of MTX. At each stage, cells were tested for their ability to differentiate in response to DMSO. RNAs were also extracted and the ratios of exogenous to

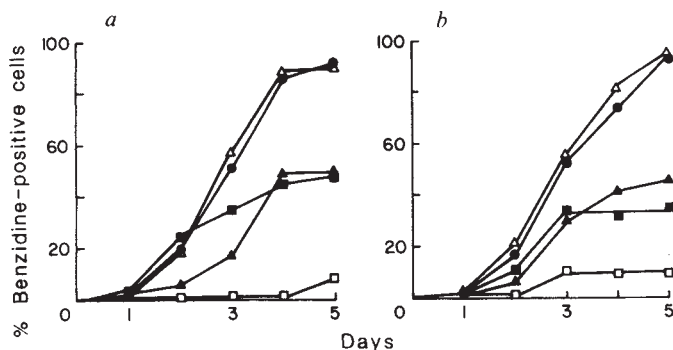


Fig. 3 Benzidine staining of F-MEL clones in response to dimethyl sulphoxide (DMSO) (a) or hypoxanthine (Hyp) (b). Equivalent numbers of cells from each of the indicated clones were cultured in the presence of 1.5% DMSO or 5 mM Hyp. At the times indicated, aliquots of cells were removed and stained with benzidine as described previously⁵. Clones 8 (■), 15 (□) and 21 (▲) were each derived from single-cell clones of F-MEL clones which had been transfected with pSV₂myc-dhfr. Clone 101 (△) was derived from a transfection with the above plasmid minus the mouse *c-myc* sequences. Clone 745 cells (●) were the non-transfected parental line used for all transfection studies. Each point represents the average of three to five independent experiments.

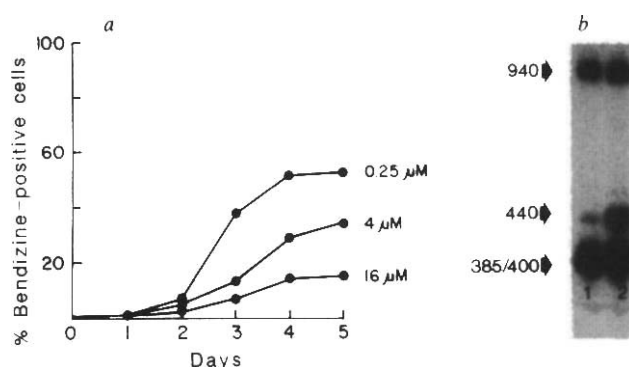


Fig. 4 The degree of F-MEL differentiation is related to the level of exogenous *c-myc* expression. Clone 21 cells were initially grown in 0.25 μM MTX. To amplify pSV₂myc-dhfr sequences, MTX concentrations were increased fourfold for 14-day intervals. At the end of the incubation period, MTX was removed for 3–5 days to allow for recovery of viable cells. Cells were then induced with 1.5% DMSO to monitor benzidine staining (a) or were subjected to the next round of MTX selection. b, After growth in the indicated concentrations of MTX, cellular RNAs were extracted and the levels of exogenous and endogenous *c-myc* transcripts were determined as described for Fig. 2. The concentrations of MTX in which clone 21 cells were grown was either 0.25 μM (lane 1) or 16 μM (lane 2).

endogenous transcripts examined. As shown in Fig. 4, benzidine staining of clone 21 cells was progressively inhibited as the concentration of MTX was raised. Exogenous *c-myc* transcript levels were also increased, indicating an amplification of plasmid sequences. These observations demonstrated that, at least in the case examined, progressive refractoriness to DMSO induction could be correlated with levels of exogenous *myc* transcripts.

The highly conserved nature of cellular oncogenes suggests that they are intimately associated with growth and/or differentiation processes. Indirect evidence in support of this comes from studies of both fresh and cultured tumour cells where aberrancies such as *c-onc* gene amplification, rearrangement and overexpression have been reported^{6–11}. Changes in the levels of transcripts of otherwise normal *c-onc* genes have been shown to accompany the normal differentiation and mitotic process both *in vitro* and *in vivo*^{12–15}.

c-myc is a short-lived nuclear protein possessing DNA-binding properties¹⁶ and is structurally related to the adenovirus E1A gene product which acts to *trans-activate* other viral genes¹⁷. Indeed, *c-myc* itself can enhance transcription of a chimaeric gene containing the 5'-flanking region of the *Drosophila hsp70* heat shock protein gene¹⁸. Together with the findings of *c-myc* transcript modulation during differentiation^{3,6}, these attributes make the *c-myc* gene product an ideal candidate to control developmental events, presumably by positively or negatively influencing the transcription of other cellular genes¹⁹. This in turn might exert effects on proliferation since, in the case of F-MEL cells, proliferation and terminal differentiation have been shown to be intimately related²⁰.

The aberrant expression of several *c-onc* genes can stimulate or inhibit differentiation in several *in vitro* systems^{21–23}. We are currently investigating the role of other *c-onc* genes in the differentiation of F-MEL cells. A more difficult problem will be that of the possible role of Friend virus complex genes in complementing the action of *c-myc* in this system.

c-myc transcripts originating from the SV40 promoter were less responsive to inducing agents than were endogenous *c-myc* transcripts. This is not surprising because the exogenous *c-myc* sequences are under the control of a foreign promoter, and the *c-myc* transcription unit has been substantially altered so as to contain a modification of the 5'-untranslated region and to lack first intron sequences. Both regions have been implicated in contributing to the considerable instability of normal *myc* mRNA²⁴. We also note that even high-level overexpression of the exogenous *c-myc* sequences did not affect the levels of endogenous transcripts. This result is consistent with previous observations regarding *c-myc* overexpression in NIH/3T3 and F-MEL cells^{25,26}.

After submission of this manuscript, Coppola and Cole reported findings similar to ours²⁶, in addition to demonstrating that some of the events leading to commitment have already occurred in *c-myc*-transfected cells.

This work was supported by grants from NIH, the Children's Leukemia Foundation of Michigan, the University of Michigan Cancer Research Institute and Strokes Against Cancer. We thank M. Imperiale for pSV₂neo, K. Marcu for pMc-myc 54 and F. Collins for use of electroporation facilities.

Received 14 April; accepted 21 May 1986.

- Marks, P. O. & Rifkind, R. A. *Rev. Biochem.* **47**, 419–448 (1978).
- Friend, C., Scher, W., Holland, J. G. & Sato, T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–382 (1971).
- Lachman, H. M. & Skoultschi, A. I. *Nature* **310**, 592–594 (1984).
- Simonsen, C. C. & Levinson, A. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2495–2499 (1983).
- Orkin, S. H., Harosi, F. I. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 98–102 (1975).
- Collins, S. & Groudine, M. *Nature* **298**, 679–681 (1982).
- Nepveu, A., Fahrlander, P. D., Yang, J.-Q. & Marcu, K. B. *Nature* **319**, 440–443 (1985).
- Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. *Science* **224**, 1121–1124 (1984).
- Rechavi, G., Givol, D. & Canaani, E. *Nature* **300**, 607–611 (1982).
- Heisterkamp, N. *et al.* *Nature* **306**, 239–242 (1983).
- Leder, P. *et al.* *Science* **222**, 765–771 (1983).
- Muller, R., Curran, T., Muller, D. & Guilbert, L. *Nature* **314**, 546–548 (1985).
- Gonda, T. J. & Metcalf, D. *Nature* **310**, 249–251 (1984).
- Slamon, D. J. & Cline, M. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7141–7145 (1984).
- Greenberg, M. E. & Ziff, E. B. *Nature* **311**, 433–438 (1984).
- Persson, H. & Leder, P. *Science* **225**, 718–721 (1984).

- Ralston, R. & Bishop, J. M. *Nature* **306**, 803–806 (1983).
- Kingston, R. E., Baldwin, A. S. & Sharp, P. A. *Nature* **312**, 280–282 (1984).
- Kingston, R. E., Baldwin, A. S. & Sharp, P. A. *Cell* **41**, 3–5 (1985).
- Gusella, J., Geller, R., Clarke, B., Weeks, V. & Housman, D. *Cell* **9**, 221–229 (1976).
- Falcone, G., Tato, F. & Alema, S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 426–430 (1985).
- Noda, M. *et al.* *Nature* **318**, 73–75 (1985).
- Waneck, G. L., Keyes, L. & Rosenberg, N. *Cell* **44**, 337–344 (1986).
- Piechaczyk, M., Yang, J.-Q., Blanchard, J.-M., Jenteur, P. & Marcu, K. B. *Cell* **42**, 589–597 (1985).
- Keath, E. J., Caimi, P. G. & Cole, M. D. *Cell* **39**, 339–348 (1984).
- Coppola, J. A. & Cole, M. D. *Nature* **320**, 760–763 (1986).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Southern, P. J. & Berg, P. J. *Molec. appl. Genet.* **1**, 327–341 (1982).
- Stanton, L. W., Fahrlander, P. D., Tesser, P. M. & Marcu, K. B. *Nature* **310**, 423–425 (1984).
- Batley, J. *et al.* *Cell* **34**, 779–787 (1983).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161–7165 (1984).
- Prochownik, E. V. *Nature* **316**, 845–848 (1985).
- Prochownik, E. V. & Orkin, S. H. *J. biol. Chem.* **259**, 15386–15392 (1984).
- Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* **65**, 718 (1980).

Vibrational CD of biopolymers

from T. A. Keiderling

Circular dichroism measurements in the infrared provide new insight into polypeptide structure and promise more detailed analyses of protein secondary structure.

Circular dichroism (CD), the differential absorption (ΔA) of left- and right-circularly polarized light, has become a fundamental tool in the analysis of solution-phase protein and nucleic acid secondary structure. The method, however, has several limitations. Ordinarily studied in the ultraviolet (UV), electronic CD can access only a few excited states of the typical organic molecule. And frequently these transitions, largely dominated by π -excitations of aromatics and carbonyls, give featureless, overlapping spectral bands whose CD tends to interfere.

On the other hand, the vibrational transitions observed in the infrared (IR) region of the spectrum are typically much better resolved and more straightforward to assign and interpret than electronic transitions. Extensive spectroscopic studies have demonstrated that some aspects of vibrational spectra are characteristic of expected biomolecular structures. But sensitivity to molecular conformation is marginal.

Coupling IR and CD

Vibrational circular dichroism (VCD) combines the details of IR data with the conformational sensitivity and two-signed clarity of CD. Thus, the variety of possible spectral transitions increases dramatically, and such transitions can originate in any bond of the molecule. The great potential of VCD lies in its ability to probe discrete structural components of a molecule, yielding spectra that are unique and very strongly influenced by conformation. Thus VCD studies complement and extend the structural insights provided by the wealth of CD and vibrational spectroscopy data already available.

VCD spectra are measured on an infrared spectrometer (conventional or FTIR) that has been modified to produce and detect modulated circularly polarized light¹. A wire grid polarizer creates linearly polarized light, which is subsequently oscillated between left and right circularity by means of a CaF₂ or ZnSe photoelastic modulator. If the sample is optically active, its VCD will in turn modulate the intensity of the incident light beam at the modulator frequency. Fast, high-sensitivity IR detectors, such as Hg(Cd)-Te or InSb, are necessary to detect the modulated intensity. The signal is extracted with lock-in amplifiers and computer signal averaging.

With this apparatus, we routinely achieve a sensitivity of $\Delta A/A$ about 10^{-5}

and a range to about 900 cm^{-1} . However, extension to approximately 600 cm^{-1} and $\Delta A/A$ of 5×10^{-5} is possible. While these capabilities for the VCD of small molecules have been available for some time, only in the last two years have applications to biopolymers developed²⁻⁴. Part of this lag can be attributed to the difficulty of studying aqueous biological systems in the infrared. For such studies, VCD must be obtained on very thin, concentrated, deuterium-exchanged samples in D₂O held between two CaF₂ windows separated by a Teflon spacer. Even then, only the amide I band (C=O stretch) can be studied with present instrumentation. In non-aqueous environments, we can also study amide A (NH stretch) and II (CNH deformation) bands^{5,6}.

Biopolymer applications

Figure 1 shows the amide I VCD and absorption spectra for three basic structural types of poly-L-lysine: random coil (rc), anti-parallel β -sheet and right-handed α -helix⁷. Both the α -helical and rc amide I bands exhibit VCD couplets at about $1,650\text{ cm}^{-1}$, but they are of opposite sense. The β -sheet form has two negative VCD bands correlating to two absorption maxima near $1,620$ and $1,680\text{ cm}^{-1}$.

In comparison to near-UV CD findings, the rc VCD results have a surprisingly large magnitude and are quite similar to the amide I VCD of left-handed α -helix previously reported⁸. This data has recently been interpreted as added evidence for substantial local order in the otherwise random coil⁹. The rc amide I VCD also resembles that of the poly-L-proline extended helix, which indicates the type of local structure present.

Poly-L-tyrosine in dimethylsulphoxide

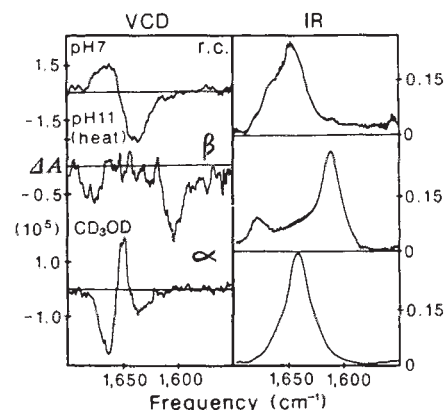


Fig. 1. Poly-L-lysine VCD and absorption spectra in the amide I region at pH 7.3 (random coil), pH 11.5 after heating (antiparallel β -sheet) and in a 96-to-4 ratio of CD₃OD to D₂O (right-handed α -helix). ΔA in units of 10^{-5} .

gives similar rc spectra. Here the amide A band can be used to distinguish the rc spectrum from that of a left-handed α -helix. Whereas interference from aromatic side chains makes electronic CD inconclusive as to secondary structure, the resolution afforded by vibrational spectra makes such analysis possible.

The VCD data set for poly-L-lysine is analogous to that used initially for characterizing protein secondary structure in electronic CD. While our present data base is insufficient to make a significant impact upon such characterizations, we have obtained globular protein VCD that reflects the qualitative patterns shown in Fig. 1. For example, the amide I VCD of myoglobin and α -chymotrypsin at neutral pH demonstrate this qualitative resemblance (Fig. 2). The predominantly α -helical myoglobin clearly shows a triple-peaked VCD of the same sign pat-

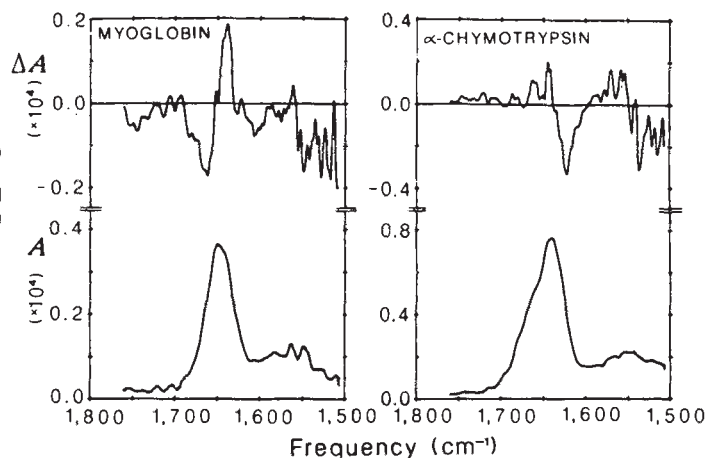


Fig. 2. Amide I VCD and absorption spectra of myoglobin and α -chymotrypsin in D₂O at neutral pH.

tern, though of smaller magnitude, as the α -helical spectrum in Fig. 1. The α -chymotrypsin, having a high β -sheet component, has a negative band at about $1,630\text{ cm}^{-1}$ that parallels the high pH poly-L-lysine results. This qualitative success suggests a role for VCD in protein stereochemical studies if improved signal-to-noise ratios and VCD for other amide bands can be obtained.

But VCD's ultimate utility in protein secondary structure analysis depends in part on its sensitivity to polymer length and end effects, which has yet to be determined. In this regard, we have made studies of β -sheet-forming¹² and 3₁₀-helical forming¹³ oligopeptides. In particular, for the latter studies, we measured the VCD of the amide A, I and II bands of aminoisobutyric acid oligopeptides as a function of length to determine when the characteristic 3₁₀-helical VCD developed. For all three bands the characteristic magnitudes of $\Delta A/A$ are fully developed at the hexamer, which corresponds to almost two turns (three H-bonds) of an ideal 3₁₀-helix. In fact the spectra evidence the same qualitative VCD as early as the tetramer, which can only form a single turn. Such results imply a dominance of relatively local interactions in VCD that is not seen in electronic CD. Thus from vibrational CD, a new level of structural insight may evolve.

Biological VCD applications are likely to focus next on the simultaneous empirical analysis of vacuum UV, near-UV and IR CD data for a series of proteins.¹⁴ Developments with FTIR-based spectrometers should also permit better signal averaging and improved resolution.¹⁵ Finally, polarized interferometers¹⁶ or new modulator designs should bring the VCD of lower energy modes within range, disclosing folded biopolymer structure in even greater detail. □

Tim Keiderling is at the Department of Chemistry, University of Illinois at Chicago, PO Box 4348, Chicago, Illinois 60680, USA.

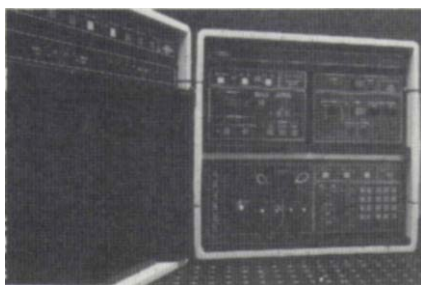
1. Keiderling, T.A. *Appl. Spectrosc. Rev.* **17**, 189-226 (1981).
2. Singh, R.D. & Keiderling, T.A. *Biopolymers* **20**, 237-240 (1981).
3. Lal, B.B. & Natie, L.A. *Biopolymers* **21**, 2161-2180 (1982).
4. Sen, A.C. & Keiderling, T.A. *Biopolymers* **23**, 1519-1532 (1984).
5. Yasui, S.C. & Keiderling, T.A. *J. Amer. Chem. Soc.* **108**, 5576-5581 (1986).
6. Paterlini, M.G., Freedman, T.B. & Natie, L.A. *Biopolymers* (in the press).
7. Yasui, S.C. & Keiderling, T.A. *Biopolymers* **25**, 5-15 (1986).
8. Narayanan, U., Keiderling, T.A., Bonora, G.M. & Toniolo, C. *Biopolymers* **24**, 1257-1263 (1985).
9. Narayanan, U., Keiderling, T.A., Bonora, G.M. & Toniolo, C. *J. Amer. Chem. Soc.* **108**, 2431-2437 (1986).
10. Yasui, S.C., Keiderling, T.A., Bonora, G.M. & Toniolo, C. *Biopolymers* **25**, 79-89 (1986).
11. Yasui, S.C., Keiderling, T.A., Formaggio, F., Bonora, G.M. & Toniolo, C. *J. Amer. Chem. Soc.* **108**, 4988-4993 (1986).
12. Hennessey, J.P. & Johnson, W.C. *Biopolymers* **20**, 1085-1094 (1981).
13. Lipp, E.D. & Natie, L.A. *Biopolymers* **24**, 799-812 (1985).
14. Dignam, M.J. & Baker, M.D. *Appl. Spectrosc.* **35**, 186-193 (1981).

ACS in Anaheim

Next week over 200 exhibitors at the American Chemical Society's West Coast meeting will be ushering in autumn with labware, instrumentation and schemes for automation.

An infrared detector built specifically for capillary gas chromatography will put in an appearance at Hewlett-Packard's stand (Reader Service No. 100). The HP 5965A system, which starts at \$47,000 (US), consists of a detection module and an HP IRD ChemStation that can combine data from both GC/FTIR and GC/MS for compound identifications. The ChemStation can also integrate library search results and rank probable internal diameters based on hit quality and common matches. Hewlett-Packard, showing in booths 141 through 147 (odd) at ACS, says the IR detector takes up less than 8 inches of bench space. The company thinks its dedicated instrument will have a heightened sensitivity for capillary GC eluents, allowing, for example, detection of 5 ng isobutyl methacrylate at a signal-to-noise ratio of 20 to 1.

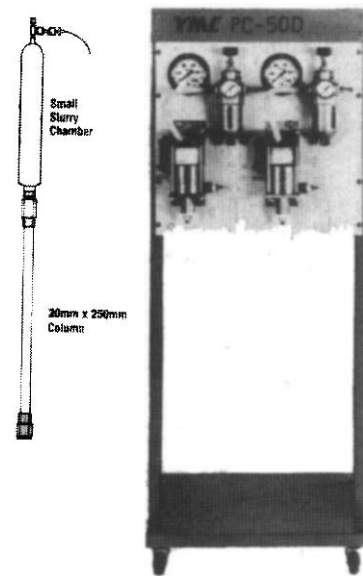
Combining non-metallic components with a novel detector array, Dionex has developed an HPLC system aimed at life sciences applications (Reader Service No. 101). The Series 4000i BioLC package has a quaternary pump that, Dionex says, can accommodate the "entire range" of HPLC techniques, including hydrophobic



The 4000i can store up to 10 programs.

interaction and ion pairing. The metal-free pump provides pressures up to 4,000 p.s.i. and flow rates from 100 μ l per minute to 10 ml per min. The 4000i detectors can spot even non-chromophoric organic and inorganic ions at sub-p.p.m. levels, according to Dionex, with a pulsed amperometric detector and suppressors to reduce background and increase sensitivity. The whole assemblage will be on display in booth 717.

For chromatographers who prefer to pack their own columns, YMC Inc has two compact slurry packing systems that, the company says, can be 10 to 15 per cent more efficient than dry or tap packing methods (Reader Service No. 102). The PC-50D uses two air-actuated pressure amplifier pumps, which generate solvent flows up to 290 ml per min at pressures up



YMC packs two columns simultaneously.

to 6,000 p.s.i., to pack LC columns with internal diameters from 4.6 mm to 20 mm. The larger PC-100D is capable of packing columns with internal diameters from 4.6 mm to 100 mm, with a larger pump to deliver solvent at 1.25 l per min at pressures up to 9,000 p.s.i. and a smaller one with the same specifications as those used in the PC-50D. YMC, at booth 705 in Anaheim, will have more information on its systems, which sell for \$3,200 (US) for the PC-50D and \$7,800 (US) for the PC-100D.

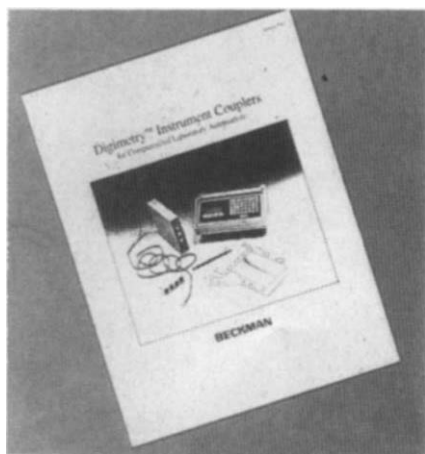
Automation links

Radiometer will have its new TitrLab wet chemistry analysis system at booth 305 in Anaheim (Reader Service No. 103). The company's programmable system automates titration procedures for both single and batch samples, switching automatically between multiple reagents, electrodes and titration procedures. Communicating via a CRT screen, TitrLab displays the titration curve during analysis and offers on-screen guidance with a HELP key. Its memory stores up to 60 user-defined methods. Radiometer's design features several different configurations for different applications to supplement the standard video titrator, triburette station and sample handler.

Software to boost the analytical prowess of Perkin-Elmer's model 7700 computer is now available for a variety of applications (Reader Service No. 104). The company claims it has written over one million lines of code to build a software selection that

spans all of its analytical instrumentation. The Perkin-Elmer Solvent Optimization System, for example, creates a multi-coloured triangular graph representing the resolution of different solvent combinations. The company sells its menu-driven software for prices ranging from \$400 (US) to \$3,000 (US), or in some cases as part of the hardware packages. Perkin-Elmer will be in a string of booths between 115 and 220 by the telephones in the Anaheim convention centre.

Three **instrument couplers** for laboratories undergoing automation are the subject of a new brochure from Beckman (Reader Service No. 105). Two of Beckman's Digimetry couplers, the MK5-4CV and the MK6, convert analogue signal outputs from chromatographs into digital format. Of the two, Beckman says, the

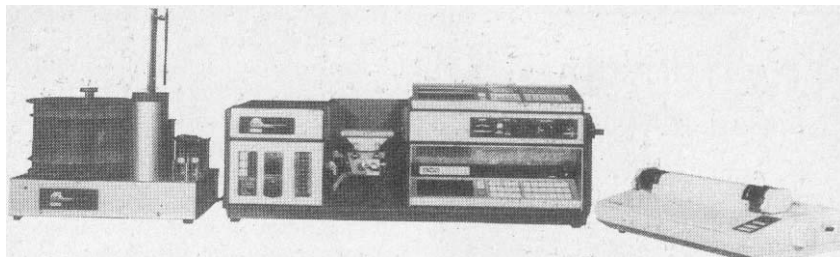


Beckman expands upon Digimetry couplers. MK5-4CV is more accurate and has on-board circuitry to validate the operational performance of each interface. The third Digimetry model, the MK5S, can collect data from as many as eight instruments with RS-232 output ports at a time. In Anaheim, Beckman will be at booths 335 through 339 (odd) to distribute further information.

Thermoanalytical determinations can now benefit from computer-aided in-



Mettler's menu-driven GraphWare system.



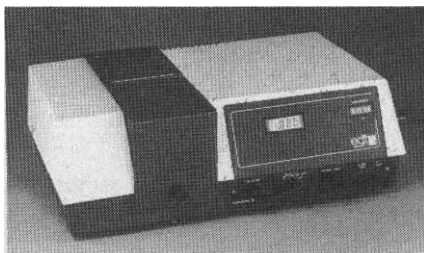
From halfway around the world, ARL's atomic absorption spectrometer has arrived.

sights, owing to new software from Mettler (Reader Service No. 106). With TA70 GraphWare, up to four thermoanalytical curves can be depicted simultaneously, either superimposed or segregated each in its own window. Mettler designed the package for use with its TA3000 thermal analysis system and an IBM PC. The company, at booth 602 in the convention hall, will have more to say about TA70 GraphWare.

Scanning the spectrum

A joint effort that spanned an ocean produced the 902 **atomic absorption spectrometer** from Applied Research Laboratories (Reader Service No. 107). Australia-based GBC Scientific Equipment, in conjunction with ARL, engineered an AA system that is guaranteed to precision of 0.5 per cent RSD on 5 p.p.m. of Cu with an absorbance of 0.7. ARL/GBC also claims that the 902's background correction system makes the instrument "the fastest system optimized for graphite furnace signals." The AA spectrometer can be found at booth 519 at ACS. A complete double-beam system with built-in data handling, software and a printer costs about \$17,000 (US).

A British-made **single beam spectrophotometer** has just arrived on the market

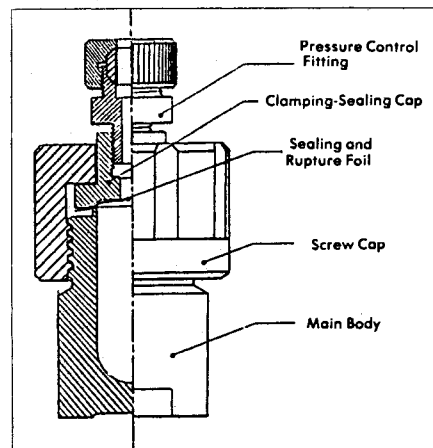


AlphaSpec: filling a gap in the UK market?

following an agreement between Camspec and EDT Research (Reader Service No. 108). The AlphaSpec range, priced between £1,500 and £2,530 (UK), is EDT's attempt to fill the gap Camspec identified for a low-cost single-beam instrument on the British market. Five AlphaSpec models are presently available, with visual or UV/vis ranges and digital or analogue readouts. EDT signed an original equipment manufacturer agreement with Camspec early this month for the AlphaSpec line, which will be distributed by the EDT group company Courtcloud, Courtcloud dealers and EDT itself.

Catalogue chronicle

Berghof/America, striving to be Teflon's titan, has a 16-page **new products cata-**



Berghof/America unveils its Teflon line.

ADVERTISEMENT



High Quality Products from FUNAKOSHI

Recombinant Mouse-Interferon-β

This is purified from E. coli carrying mouse IFN-β gene of mouse L-cell, using affinity chromatographies described by S. Matsuda, et al. (to be published in J. Interferon Res).

Purity: higher than 99% by SDS-PAGE.



Blue-Cotton (Mutagen Adsorbent)

This is able to adsorb mutagens containing quite low concentration in aqueous solutions, especially those having more than three fused aromatics.

The adsorbed mutagens are easily eluted by ammonia-methanol solution.

Blue-cotton is applied to Urine, Feces, Foods, River Water, etc.



FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone: Tokyo 03-293-2352
Telefax: 81-3-295-5545
Telex: J28489 FUNA

ADVERTISEMENTS

LIMUTESTER

(gelation-capillary method)



Precise Detection of Endotoxins

This LIMUTESTER detects Endotoxins promptly, applying to gelation reaction of LAL and a glass capillary. The LIMUTESTER is simple, economical and usable anywhere, especially needs no endotoxin-free laboratory glass wares. LIMUTESTER INCUBATOR is available.

FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone : Tokyo 03-293-2352
Telefax : 81-3-295-5545
Telex : J28489 FUNA

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low
Cost



Eliminates evaporation loss and contamination. Easy to use — flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.

Medwire Div.

stocks a wide variety of

TEFLON COATED*

Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.



Medwire Div.

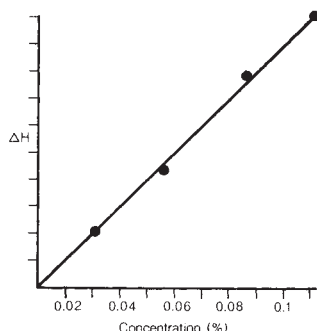
Affiliate of Sigmund Cohn Corp.
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300

*Reg. T.M. of E.I. DuPont

Reader Service No.49

logue that outlines its latest laboratory equipment (Reader Service No. 109). Large volume high pressure autoclaves, Teflon-encased stainless-steel stir shafts, flexible corrugated tubing up to 150 feet long and pressure digestion heating blocks are just a few of the areas that incorporate PTFE, FEP, PFA and other fluorocarbon resins. Berghof will be showing its equipment at booth 428 in Anaheim.

The Bioanalyzer brochure from Provesta details the theory and practice of its **enzyme-electrode instrument** (Reader Service No. 110). The Bioanalyzer,

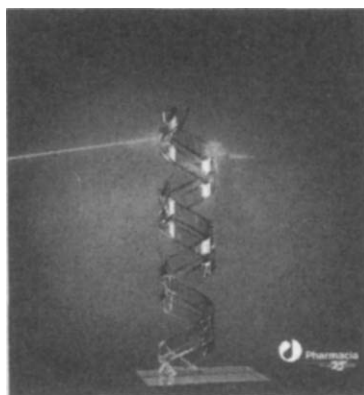


A typical standard curve for the Bioanalyzer.

according to Provesta, can determine the concentration of specific biochemicals or organic chemicals in complex mixtures at levels between 10^{-5} and 10^{-1} M. Using an oxygen electrode tipped with an oxidase in a gel matrix, the instrument works by measuring the decrease in oxygen at the electrode when the oxidase catalyses a reaction between oxygen and the biochemical being analysed. Information on Wednesday's Bioanalyzer workshop will be on hand at Provesta's ACS booth 726.

Pharmacia's catalogue reviews the company's **recent arrivals for the molecular biologist** (Reader Service No. 111). What isn't on exhibit at the biotech group's ACS booth 304 will surely be listed among the catalogue's pages. It describes many of the products Pharmacia has advanced during

Molecular Biologicals

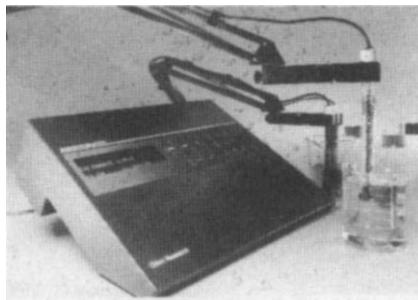


Pharmacia tells all in its new catalogue.

the past year, including its DNA/RNA modifying enzymes, prokaryotic and eukaryotic vectors, DNA synthesizer and linkers for protein engineering.

For the life scientist

Right near the convention hall entrance, Orion Research in booths 702 and 704 will be touting the EA 940 **pH/ISE meter** with its unusual ability for expansion (Reader Service No. 112). Orion says the meter alone can be programmed to measure ions for which electrode and direct calibration procedures do not exist, or to measure samples with unknown concentrations and different sample types without calibration. But by adding the company's 960 autochemistry module, the meter can also be upgraded to a system capable of performing analyses ranging from endpoint titrations to incremental techniques. The



Three keys control most EA940 functions.

EA940 sells for \$1,995 in the United States.

Ohaus will debut its high-capacity top-loading balance at ACS in booths 735 and



New heights in weights for Ohaus balances.

739 (Reader Service No. 113). The Galaxy 8000 weighs up to 8,000 g and reads to 0.1g; its taring range is also 0 to 8,000 g. The G8000's 7 x 9-inch weighing pan reflects its high capacity function. Ohaus also equipped its unit for RS-232 bi-directional data output and sample weight display in grams, pounds, troy or avoirdupois ounces, carats or pennyweights. Taring and calibration, says Ohaus, are accomplished with the touch of a button. The G8000 lists at \$1,150 (US). □

These notes are compiled by Karen Wright from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.